APPLICATION

FOR

UNITED STATES LETTERS PATENT

TITLE: INTRAFLAGELLAR TRANSPORT

APPLICANT: GEORGE B. WITMAN, GREGORY J. PAZOUR, JOEL L. ROSENBAUM AND DOUGLAS G. COLE

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL298428026

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

May 24, 2001
Date of Deposit

Signature

Lisa G. Gray
Typed or Printed Name of Person Signing Certificate

INTRAFLAGELLAR TRANSPORT

Cross Reference To Related Applications

This application claims priority from U.S. Provisional Application Serial No. 60/206,923, filed on May 24, 2000, which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

This invention relates to intraflagellar transport proteins.

BACKGROUND

Cilia are tiny cellular structures that protrude from cells. They are about 0.25 micrometers in diameter and contain a bundle of microtubules. They are widespread among living organisms, occurring in most animals, many single-celled eukaryotes, and in some lower plants.

Cilia tend to function in one of two ways. They may move fluid across the surface of the cell or they may propel cells through a fluid. They may also serve to gather food. In humans, cilia on the surfaces of respiratory epithelia function to push mucus and trapped particles and dead cells out of the lungs. Cilia also function to carry eggs through the oviduct. Cilia function in myriad ways in different kinds of cells.

Flagella are structures related to cilia. They are similar in internal structure but tend to be much longer than cilia. Sperm cells are propelled by flagella, as are many other single-celled eukaryotes.

Groups of cilia tend to move together in coordinated unidirectional waves. The motion made by each individual cilium is whiplike. This motion includes two phases. First, the cilium extends forward, pushing against the surrounding liquid as it goes. At the end of its forward stroke, the cilium bends, reducing viscous drag as it pulls itself back to its original position. By contrast, flagella tend to propagate quasi-sinusoidal waves. Despite the differences in their external motions, the molecular basis of movement in both cilia and flagella appear to be the same.

Cilia and flagella move by bending their core – the axoneme. The axoneme is composed of microtubules and associated proteins. The pattern of microtubules is distinctive: nine pairs of microtubules that form a ring around two single microtubules. This arrangement is typically referred to as "9 + 2". The pairs are composed of one complete and one partial microtubule. These microtubules extend the full length of the axoneme, which can range in length from 10-200 micrometers.

Intraflagellar transport (IFT) is a dynein and kinesin-based motility process in which non-membrane-bound particles move along flagellar microtubules, just beneath the flagellar membrane, from the base to the tip of the flagellum and back. IFT is essential for the assembly and maintenance of all cilia and flagella, including non-motile primary cilia and sensory cilia. Recent results indicate that defects in IFT are a primary cause of several human diseases.

SUMMARY

The invention is based on the discovery of various new IFT particle polypeptides and the genes that encode them.

In general, the invention features, an isolated nucleic acid molecule selected from the group consisting of: a) a nucleic acid molecule having a nucleotide sequence which is at least 90% identical to the nucleotide sequence of Chlamydomonas intraflagellar transport (IFT) particle protein gene 20, 27, 46, 52, 57, 72, 88, 122, 139, or Che-2, or a complement thereof; b) a nucleic acid molecule comprising at least 15 nucleotide residues and having a nucleotide sequence identical to at least 15 consecutive nucleotide residues of the nucleotide sequence of Chlamydomonas IFT particle protein gene 20, 27, 46, 52, 57, 72, 88, 122, or 139, or Che-2, or a complement thereof; c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of Chlamydomonas IFT particle protein 20, 27, 46, 52, 57, 72, 88, 122, 139, or Che-2; or d) a nucleic acid molecule which encodes a polypeptide comprising at least 10 amino acids and having an amino acid sequence identical to at least 10 consecutive amino acids of the amino acid sequence of Chlamydomonas IFT particle protein 20, 27, 46, 52, 57, 72, 88, 122, 139, or Che-2.

The nucleic acid molecules can further include nucleic acid sequences encoding a heterologous polypeptide. The invention also features a vector including the nucleic acid molecules and host cells including the new nucleic acid molecules, such as non-human mammalian host cells.

The invention also features an isolated polypeptide selected from the group consisting of: a) a polypeptide comprising at least 10 amino acids and having an amino acid sequence identical to at least 10 consecutive amino acids of the amino acid sequence of Chlamydomonas intraflagellar transport (IFT) particle protein 20, 27, 46, 52, 57, 72, 88, 122, 139, or Che-2; b) a polypeptide comprising the amino acid sequence of Chlamydomonas IFT particle protein 20, 27, 46, 52, 57, 72, 88, 122, 139, or Che-2, wherein the polypeptide comprises one or more conservative amino acid substitutions that do not inhibit the biological activity of the polypeptide relative to a corresponding native Chlamydomonas IFT particle protein; and c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 90% identical to a nucleic acid consisting of the nucleotide sequence of Chlamydomonas IFT particle protein gene 20, 27, 46, 52, 57, 72, 88, 122, 139, or Che-2, or a complement thereof.

In another aspect, the invention features an antibody that selectively binds to the new polypeptides.

Yet other nucleic acid molecules of the invention include those having sequences that (1) are at least 90% identical to the nucleotide sequence of mouse intraflagellar transport (IFT) particle protein gene 57 (or are complements thereof); (2) are at least 15 nucleotide residues long and have a sequence identical to at least 15 consecutive nucleotide residues of the nucleotide sequence of mouse IFT particle protein gene 57 (or complements thereof); (3) encode a polypeptide that is or that includes the amino acid sequence of mouse IFT particle protein 57; or (4) encode a polypeptide having at least 10 amino acids and an amino acid sequence identical to at least 10 consecutive amino acids of the amino acid sequence of mouse IFT particle protein 57. For example, a nucleic acid molecule of the invention can be a nucleic acid having the nucleotide sequence of mouse IFT particle protein gene 57 (or a complement thereof) or a nucleic acid molecule that encodes a polypeptide having the amino acid sequence of mouse IFT particle protein 57.

Yet other polypeptides of the invention include those having sequences that (1) include at least 10 amino acid residues and have an amino acid sequence identical to at least 10 consecutive amino acids of the amino acid sequence of mouse intraflagellar transport (IFT) particle protein 57; (2) include the amino acid sequence of mouse IFT particle protein 57; or (3) are encoded by a nucleic acid molecule having a nucleotide sequence that is at least 90% identical to a nucleic acid consisting of the nucleotide sequence of mouse IFT particle protein gene 57 (or a complement thereof). For example, a polypeptide of the invention can include the amino acid sequence of mouse

IFT particle protein 57. Any of the polypeptides described herein can include one or more conservative amino acid substitutions that do not inhibit the biological activity of the polypeptide relative to native mouse IFT particle protein 57 (*e.g.*, polypeptides that retain at least 50% (*e.g.* 60%, 75%, 80%, 90%, or 95% or more) of one or more of the biological activities of a native IFT polypeptide (*e.g.* particle protein 57)).

Various methods are also within the scope of the invention. For example, the invention features a method for identifying a candidate compound that modulates (*e.g.*, inhibits or stimulates) the activity of mouse intraflagellar transport (IFT) particle protein 57. The method can be carried out, for example, by contacting a test compound with an isolated IFT particle polypeptide and determining whether the test compound interacts with the polypeptide. Interaction indicates that the test compound is a candidate modulator of mouse IFT particle protein 57. Similarly, one can carry out methods to identify a candidate compound that modulates (*e.g.*, inhibits or stimulates) the activity of a human intraflagellar transport (IFT) particle protein. These methods can be carried out, for example, by contacting a test compound with an isolated IFT particle polypeptide and determining whether the test compound interacts with the polypeptide. Here again, interaction indicates that the test compound is a candidate modulator of a human IFT particle protein. The isolated human IFT particle polypeptide can be human IFT particle polypeptide 20-1, 20-2, 20-3, 27, 46, 52, 57-1, 57-2, 72, 88, 122, 139-1, 139-2 or Che-2.

The methods described above can include additional steps. For example, one can contact that candidate modulator with one or more cells (*e.g.* cultured cells) that have functional cilia and determine whether the modulator modulates (*e.g.* inhibits or stimulates) cilia function. In the event cilia function is inhibited, the candidate modulate is an IFT particle protein inhibitory agent. In other embodiments, the candidate modulator can be contacted with one or more cells (*e.g.* cultured cells) that have non-functional or functionally impaired cilia (*e.g.* a cell or cells lacking a specific IFT particle protein). Restoration (partial or complete) of cilia function indicates that the candidate modulator is an IFT particle protein restorative agent.

Other methods of the invention can be used to identify a candidate compound that restores the activity of a defective or absent human intraflagellar transport (IFT) particle protein. These methods can be carried out by, for example, obtaining a mixture of isolated IFT particle polypeptides that include (i) all but one of the IFT particle polypeptides required to form the IFT particle, and (ii) a medium that enables the IFT particle polypeptides to form the IFT particle when all normal IFT

particle polypeptides that constitute that IFT particle are present. The mixture is then contacted with a test compound, after which one determines whether the test compound enables the IFT particle to be formed. IFT particle formation indicates that the test compound is a candidate compound that restores the activity of a defective or absent human IFT particle protein. These methods can be carried out by contacting the candidate compound with one or more cells (*e.g.*, cultured cells) that have non-functional (or impaired) cilia and that lack a specific IFT particle protein. One can then determine whether the candidate compound restores cilia function, which would indicate that the candidate compound is an IFT particle protein restorative agent. These methods (and others of the invention) can be carried out with a human IFT particle polypeptide (*e.g.*, human IFT particle polypeptides 20-1, 20-2, 20-3, 27, 46, 52, 57-1, 57-2, 72, 88, 122, 139-1, 139-2 or Che-2; or combinations thereof).

Diagnostic methods are also within the scope of the invention. For example, the invention features a method of diagnosing a disorder in a tissue in a subject that is associated with (*e.g.* caused by, in whole or in part) a defective or absent human intraflagellar transport (IFT) particle protein. The method can be carried out, for example, by disrupting cells from a tissue sample, contacting the disrupted cells with an antibody that specifically binds to a normal human IFT particle protein, and detecting binding of the antibody to any IFT particle protein in the sample. An absence of specific binding indicates that the tissue is one in which IFT particle protein is defective or absent. Disorders (or diseases or conditions) associated with this tissue defect include, but are not limited to, kidney disease, retinal disorders, thyroid disorders, chondrocyte disease, olfactory disease, azoospermia, and primary ciliary dyskinesia.

In other embodiments, the methods are methods of treatment. For example, the invention features a method of treating a disorder in a subject (*e.g.* a disorder associated with a defective or absent intraflagellar transport (IFT) protein) by administering to the subject a human IFT particle polypeptide. The amount of the polypeptide is an amount that is effective to compensate for the defective or absent IFT particle protein. Rather than administering a polypeptide of the invention, one can administer a nucleic acid molecule that encodes it. For example, one can administer a human IFT particle polypeptide (*e.g.*, human IFT particle polypeptides 20-1, 20-2, 20-3, 27, 46, 52, 57-1, 57-2, 72, 88, 122, 139-1, 139-2 or Che-2) or a nucleic acid molecule encoding such a polypeptide. The polypeptides or nucleic acid molecules can be delivered with a pharmaceutically acceptable carrier, excipient, or diluent.

The methods of the invention can also be carried out to treat an infection in a subject that is caused by a pathogen (*e.g.* a nematode, bacteria, protozoa or insect) that has an intraflagellar transport (IFT) particle protein. The subjects are treated by administering to them an effective amount of an agent that inhibits the function of the IFT particle protein (*e.g.*, an antibody that binds specifically to the IFT particle protein). The diagnostic and therapeutic methods of the invention can be used to diagnose or treat mammals (*e.g.* humans or other primates, or domesticated or farm animals) as well as plants.

Another suitable method for identifying compounds that inhibit or restore IFT function involves screening for small molecules that specifically bind to IFT proteins. A variety of suitable binding assays are known in the art as described, for example, in U.S. Patent Nos. 5,585,277 and 5,679,582, incorporated herein by reference. For example, in various conventional assays, test compounds can be assayed for their ability to bind to a polypeptide by measuring the ability of the small molecule to stabilize the polypeptide in its folded, rather than unfolded, state. More specifically, one can measure the degree of protection against unfolding that is afforded by the test compound. Test compounds that bind to an IFT protein with high affinity cause, for example, a significant shift in the temperature at which the polypeptide is denatured. Test compounds that stabilize the polypeptide in a folded state can be further tested for IFT inhibitive or restorative activity in a standard susceptibility assay.

The IFT polypeptides also can be used in assays to identify test compounds that bind to the polypeptides. Test compounds that bind to IFT polypeptides then can be tested, in conventional assays, for their ability to inhibit or restore IFT function. Test compounds that bind to IFT polypeptides are candidate IFT inhibitive or restorative agents, in contrast to compounds that do not bind to IFT polypeptides. As described herein, any of a variety of art-known methods can be used to assay for binding of test compounds to IFT polypeptides. If desired, the test compound can be immobilized on a substrate, and binding of the test compound to an IFT polypeptide is detected as immobilization of an IFT polypeptide on the immobilized test compound. Immobilization of an IFT polypeptide on the test compound can be detected in an immunoassay with an antibody that specifically binds to an IFT polypeptide.

Also included in the invention is a method for identifying a candidate IFT restorative agent useful for treating abnormal IFT function by: (a) contacting a polypeptide encoded by an IFT nucleic acid with a test compound; and (b) detecting binding of the test compound to the

polypeptide, wherein a compound that binds to the IFT polypeptide indicates that the compound is a candidate IFT function restorative agent, and wherein the polypeptide is encoded by a gene selected from the group consisting of a first nucleic acid molecule which encodes a polypeptide containing the amino acid sequence of a polypeptide of the invention or a naturally occurring allelic variant of a polypeptide of the invention, wherein the first nucleic acid molecule hybridizes to a second nucleic acid molecule under stringent conditions. The method can further include a step of determining whether the candidate compound that binds to the IFT polypeptide inhibits growth of cells or organisms, relative to growth of cells or organisms grown in the absence of the test compound, wherein inhibition of growth indicates that the candidate compound is an anti-IFT agent.

In one example, the test compound is immobilized on a substrate, and binding of the test compound to the IFT polypeptide is detected as immobilization of the IFT polypeptide on the immobilized test compound. Immobilization of the IFT polypeptide on the test compound can be detected in an immunoassay with an antibody that specifically binds to the IFT polypeptide.

In one example, the test compound is selected from the group consisting of polypeptides, ribonucleic acids, small molecules (e.g., organic or inorganic), aptamers, peptidomimetics, carbohydrates, and deoxyribonucleic acids.

By "operably linked" is meant that a gene and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are bound to the regulatory sequence(s).

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

DESCRIPTION OF DRAWINGS

FIG. 1 is a representation of a model of the IFT particles involved in intraflagellar transport towards and away from the tip of a flagellum.

FIG. 2 is a representation of a large scale model of the IFT particles involved in intraflagellar transport towards and away from the tip of the flagellum.

FIG. 3 is a representation of the flagellar pore complex.

FIG. 4 is a representation of a model of how flagellar proteins are targeted to the flagellum.

FIG. 5 is a representation of a model of the photoreceptor cell IFT.

FIG. 6A is a representation of *Chlamydomonas* IFT20 particle polypeptide (SEQ ID NO:2).

FIG 6B is a representation of *Chlamydomonas* IFT20 particle nucleic acid (SEQ ID NO:1).

FIG. 6C is a representation of human IFT20 particle polypeptide (SEQ ID NO:23).

FIG. 6D is a representation of human IFT20 particle polypeptide (SEQ ID NO:24).

FIG. 6E is a representation of human IFT20 particle polypeptide (SEQ ID NO:25).

FIG. 7A is a representation of *Chlamydomonas* IFT27 particle polypeptide (SEQ ID NO:4).

FIG 7B is a representation of *Chlamydomonas* IFT27 particle nucleic acid (SEQ ID NO:3).

FIG. 7C is a representation of human IFT27 particle polypeptide (SEQ ID NO:26).

FIG. 8A is a representation of *Chlamydomonas* IFT46 particle polypeptide (SEQ ID NO:6).

FIG 8B is a representation of *Chlamydomonas* IFT46 particle nucleic acid (SEQ ID NO:5).

FIG. 8C is a representation of human IFT46 particle polypeptide (SEQ ID NO:27).

FIG. 9A is a representation of *Chlamydomonas* IFT52 particle polypeptide (SEQ ID NO:8).

FIG 9B is a representation of *Chlamydomonas* IFT52 particle nucleic acid (SEQ ID NO:7).

FIG. 9C is a representation of human IFT52 particle polypeptide (SEQ ID NO:28).

FIG. 9D is a representation of *Caenorhabditis elegans* IFT52 particle polypeptide (SEQ ID NO:29).

FIG. 10A is a representation of *Chlamydomonas* IFT57 particle polypeptide (SEQ ID NO:10).

FIG 10B is a representation of *Chlamydomonas* IFT57 particle nucleic acid (SEQ ID NO:9).

FIG. 10C is a representation of mouse IFT57 particle polypeptide (SEQ ID NO:12).

FIG 10D is a representation of mouse IFT57 particle nucleic acid (SEQ ID NO:11).

FIG. 10E is a representation of human IFT57 particle polypeptide (SEQ ID NO:30).

FIG. 10F is a representation of human IFT57 particle polypeptide (SEQ ID NO:31).

FIG. 10G is a representation of *Caenorhabditis elegans* IFT57 particle polypeptide (SEQ ID NO:32).

FIG. 11A is a representation of *Chlamydomonas* IFT72 particle polypeptide (SEQ ID NO:14).

FIG 11B is a representation of *Chlamydomonas* IFT72 particle nucleic acid (SEQ ID NO:13).

FIG. 11C is a representation of human IFT72 particle polypeptide (SEQ ID NO:33).

FIG. 12A is a representation of *Chlamydomonas* IFT88 particle polypeptide (SEQ ID NO:16).

FIG 12B is a representation of *Chlamydomonas* IFT88 particle nucleic acid (SEQ ID NO:15).

FIG. 12C is a representation of human IFT88 particle polypeptide (SEQ ID NO:34).

FIG. 12D is a representation of *Caenorhabditis elegans* IFT88 particle polypeptide (SEQ ID NO:35).

FIG. 13A is a representation of *Chlamydomonas* IFT122 particle polypeptide (SEQ ID NO:18).

FIG 13B is a representation of *Chlamydomonas* IFT122 particle nucleic acid (SEQ ID NO:17).

FIG. 13C is a representation of human IFT122 particle polypeptide (SEQ ID NO:36).

FIG. 13D is a representation of *Caenorhabditis elegans* IFT122 particle polypeptide (SEQ ID NO:37).

FIG. 14A is a representation of *Chlamydomonas* IFT139 particle polypeptide (SEQ ID NO:20).

FIG 14B is a representation of *Chlamydomonas* IFT139 particle nucleic acid (SEQ ID NO:19).

FIG. 14C is a representation of human IFT139 particle polypeptide (SEQ ID NO:38).

FIG. 14D is a representation of human IFT139 particle polypeptide (SEQ ID NO:39).

FIG. 14E is a representation of *Caenorhabditis elegans* IFT139 particle polypeptide (SEQ ID NO:40).

FIG. 15A is a representation of *Chlamydomonas* Che-2 polypeptide (SEQ ID NO:22).

FIG 15B is a representation of *Chlamydomonas* Che-2 nucleic acid (SEQ ID NO:21).

FIG. 15C is a representation of human Che-2 polypeptide (SEQ ID NO:41).

FIG. 15D is a representation of *Caenorhabditis elegans* Che-2 polypeptide (SEQ ID NO:42).

DETAILED DESCRIPTION

A. Intraflagellar Transport

Intraflagellar transport (IFT) is a dynein and kinesin-based motility process in which non-membrane-bound particles move along flagellar microtubules, just beneath the flagellar membrane, from the base to the tip of the flagellum and back (see Fig. 1). IFT is essential for the assembly and maintenance of all cilia and flagella, including non-motile primary cilia and sensory cilia. Recent results indicate that defects in IFT are a primary cause of several human diseases.

Much of our knowledge of the protein machinery and basic biology of IFT has come from studies of the bi-flagellate alga Chlamydomonas, which is an excellent model system for biochemical and molecular genetic analyses of proteins and processes occurring in the flagellum. IFT was first observed in Chlamydomonas using Differential Interference Contrast (DIC) microscopy (Kozminski et al., Proc. Nat. Acad. Sci. U.S.A., 90:5519-5523, 1993), where particles were seen to move anterogradely (i.e., from base to tip) along the flagellum at ~2.5 mm/s without pause, while apparently smaller particles moved retrogradely at ~4 mm/s. Correlative light and electron microscopy revealed that the particles being moved during IFT

were organized into linear arrays termed "rafts" (Kozminski et al., Proc. Nat. Acad. Sci. U.S.A., 90:5519-5523, 1993). Electron microscopy of the arrays showed that they were connected by periodic links to both the flagellar membrane and the B-tubule of the outer doublet microtubule (Kozminski et al., Proc. Nat. Acad. Sci. U.S.A., 90:5519-5523, 1993; Kozminski et al., J. Cell Biol., 131:1517-1527, 1995; Pazour et al., J. Cell Biol, 141:979-992, 1998). Biochemical characterization of the Chlamydomonas IFT particles revealed that they are composed of at least 16 different polypeptides (Piperno and Mead, Proc. Natl. Acad. Sci., 94:4457-4462, 1997; Cole et al., J. Cell Biol., 141:993-1008, 1998); many of these are now known to have homologues in higher organisms, including humans. More recently, IFT has been visualized in the ciliated sensory neurons of C. elegans (Orozco et al., Nature, 398:674, 1999; Signor et al., J. Cell Biol., 147:519-530, 1999; Qin et al., Cur. Bio., 11:1-20, 2001) and in the primary cilia of mouse kidney cells (Pazour, manuscript in preparation) using GFP-tagged IFT proteins. Defects in IFT motors and particle proteins lead to defects in assembly of motile flagella in Chlamydomonas (Kozminski et al., J. Cell. Biol., 131:1517-1527, 1995; Pazour et al., J. Cell Biol., 144:473-481, 1999; Pazour et al., J. Cell Biol., 151:709-718, 2000), of ciliated sensory neurons in C. elegans (Perkins et al., *Dev. Biol.*, 117:456-487, 1986; Cole et al., J. Cell Biol., 141:993-1008, 1998; Collet et al., Genetics, 148:187-200, 1998; Signor et al., J. Cell Biol., 147:519-530, 1999; Wicks et al., Dev. Biol., 221:295-307, 2000; Qin et al., Cur. Bio., 11:1-20, 2001), and of nodal cilia (Nonaka et al., Cell, 95:829-837, 1998; Marszalek et al., Proc. Natl. Acad. Sci. U.S.A., 96:5043-5048, 1999; Takeda et al., J. Cell Biol., 145:825-836, 1999; Murcia et al., Development, 127:2347-2355, 2000), kidney primary cilia (Pazour et al., J. Cell Biol., 151:709-718, 2000), and rod outer segments in mice (Pazour et al., manuscript in preparation). IFT functions in the transport of flagellar precursors to the tip of the flagellum, where they are needed for both flagellar assembly and maintenance. In cilia and flagella with a sensory function, IFT can also be involved in signal transduction between the cilium or flagellum and the cell body.

The motors powering IFT

The motors moving the IFT particles were defined by means of Chlamydomonas flagellar mutants defective in either kinesin-II or cytoplasmic dynein 1b.

The anterograde motor. Studies of flagellar assembly (fla) mutants in Chlamydomonas first identified the gene FLA10, which encodes a kinesin-II motor subunit (termed Fla10) present

in the flagellum. IFT and Fla10 were essential for both the assembly and maintenance of the flagella. Inasmuch as kinesin-II moves towards the + ends of microtubules (Scholey, J. Cell Biol., 133:1-4, 1996), it was proposed that anterograde IFT was powered by kinesin-II (Kozminski et al., J. Cell Biol., 131:1517-1527, 1995). The flagellar Fla10-kinesin-II was purified to homogeneity and shown to be a typical heterotrimeric kinesin-II, composed of two motor subunits of 85 and 95 kD (also called Kif 3a and 3b [Scholey, J. Cell Biol., 133:1-4, 1996]), and a non-motor subunit of 115 kD, called KAP (kinesin-associated protein, [Scholey, J. Cell Biol., 133:1-4, 1996]).

The retrograde motor. Cytoplasmic dynein was first implicated in IFT when it was found that a mutation in Chlamydomonas LC8, a cytoplasmic dynein light chain, had short flagella that initially grew out to about one-half to three-quarters normal length, and then gradually shortened (Pazour et al., 1998). These flagella exhibited normal levels of anterograde IFT, but had greatly reduced levels of retrograde IFT.

The Chlamydomonas dynein heavy chain isoform DHC1b forms the retrograde motor. In C. elegans, the homologue of DHC1b is Che-3, and mutations in Che-3 result in a phenotype very similar to that seen in DHC1b mutants in Chlamydomonas: the sensory cilia are very short and filled with IFT particles (Collet et al., 1998; Wicks et al., 2000). Therefore, DHC1b (also known as DHC2 in mammals) is likely to be the retrograde IFT motor in all cilia and flagella.

IFT and the tip of the flagellum

IFT particles move all the way to the tip or base of the flagellum without pause; particles moving to the tip appear by DIC microscopy to have more contrast than those moving to the base (Kozminski et al., 1993; Piperno et al., 1998). Figure 2 shows a model of this process. Analysis of *Chlamydomonas* IFT motor mutants indicates that kinesin-II is carried as IFT cargo from the flagellar tip to the base (Pazour et al., 1998, 1999); conversely, because dyneins are minus-end directed motors, it can be inferred that cytoplasmic dynein 1b is carried as IFT cargo to the tip. There is evidence that flagellar precursor proteins also are transported by IFT particles to the site of axonemal assembly at the flagellar tip, and released there (Piperno et al., 1996). At the tip, the apparent size of the particles is reduced (possibly due to unloading of axonemal precursors), the kinesin motor becomes cargo, and the cytoplasmic dynein 1b motor takes over to transport the particles back to the peri-basal body region.

The mechanisms by which IFT particle turnaround, cargo-release, and motor exchange occur at the tip are unknown, but by analogy with other bi-directional particle movement systems, e.g. melanophore movement (Reese and Haimo, 2000), it may involve phosphorylation and dephosphorylation of motors and/or their associated proteins. Because the IFT particles move unidirectionally without stopping or reversing, the proteins that control the IFT motors must be highly localized at the tip and base of the flagellum. At the base of the flagellum, the regulatory proteins may be anchored to the transition fibers (Deane et al., manuscript in preparation). The *tip* of the flagellum also contains specialized structures that may serve as anchors for the proteins that turn kinesin off and dynein on: each of the A-subfibers of the outer doublets is terminated by a plug which is connected to the flagellar membrane by a thin filament (Dentler and Rosenbaum, 1977; Dentler, 1989). The central pair of microtubules likewise is connected to the tip by a specific structure (Dentler and Rosenbaum, 1977), but this probably is not involved in the turnaround process because *Chlamydomonas* mutants lacking the central pair have normal IFT (Kozminski et al., 1993).

Characterization of Chlamydomonas IFT particles

When fla10ts mutant cells are shifted to restrictive temperature, the number of IFT particles in the flagella decreases substantially before the flagella are resorbed (Kozminski et al., 1995). By comparing these fla10 mutant flagella to those of wild-type cells, it was possible to determine which flagellar polypeptides make up the IFT particles. In this way, the IFT particles were found to sediment at ~16 S in sucrose gradients and to be composed of at least 16 polypeptides (Piperno and Mead, 1997; Cole et al., 1998), occurring in two complexes, termed A and B. Complex A contains 4 polypeptides of relatively high molecular weight (Mr 122K-144K), whereas Complex B contains 11-12 mostly lower MW polypeptides (Mr < 100K) (Cole et al. 1998). The identity of the proteins was confirmed by analysis of LC8 mutant flagella, which accumulated both the IFT polypeptides and IFT particles (Pazour et al., 1998).

Separation of the IFT particle polypeptides on two-dimensional gels permitted microsequencing of individual proteins to obtain short internal amino-acid sequences (Cole et al., 1998). The peptide sequences were in turn used to generate PCR primers for amplification of cDNAs encoding several of the IFT particle polypeptides (see below). To date, the predicted sequences are novel. The sequences have permitted the identification of Chlamydomonas

insertional mutants lacking the genes encoding two of the IFT particle proteins, IFT57 and IFT88 (Pazour et al., 1999b; Pazour et al., 2000). These mutants grow and divide normally, demonstrating that these IFT particle proteins are not involved in any essential processes. However, the IFT88 mutant fails to assemble flagella, whereas the IFT57 mutant assembles only very short flagella. Thus, these specific IFT particle proteins are required for flagellar formation, although the difference in phenotypes suggests that the two proteins differ in their importance. In contrast to the DHC1b mutant, the IFT57 and IFT88 mutants do not accumulate IFT particles in their flagella.

Localization of IFT particle proteins

Isolation of the IFT particles permitted the production of both monoclonal and polyclonal antibodies recognizing specific IFT particle proteins. Immunofluorescence microscopy using these antibodies and those prepared against the IFT motors revealed punctate staining along the flagella, presumably representing IFT particles in transit. However, the principal localization of both the IFT particle polypeptides and the kinesin and dynein motors was in a circular pattern around the two basal bodies/centrioles (Cole et al., 1998; Pazour et al., 1999). This was somewhat surprising, because moving IFT particles had been observed only in the flagella by DIC microscopy, and the linear arrays of IFT particles had been observed only in the flagella by electron microscopy. This notwithstanding, the immunoflourescence localization pattern indicated that there is a large pool of IFT particle proteins and motors at the base of the flagellum.

To learn more about this peri-basal body distribution of IFT particle proteins, studies were carried out using gold-labeled antibodies and thin-sectioned material (Deane et al., manuscript in preparation). These higher resolution studies revealed that the IFT proteins were localized at the "flagellar" end of the basal bodies, specifically on the membrane-associated ends of the transition fibers that connect the basal body to the cell membrane (Ringo, 1967; Weiss et al., 1977). These fibers demarcate the boundary between the cytoplasmic and flagellar "compartments." Although the flagellum is ostensibly "open" to the cytoplasm, it appears that only a subset of cytoplasmic proteins (the "flagellar" proteins) gain admission to the flagellar compartment. Thus, the transition fibers may be structural components of a filter or "flagellar pore" that controls movement of molecules and particles between the cytoplasmic and flagellar

compartments, much as the nuclear pore controls movement between the cytoplasmic and nuclear compartments. If access to the flagellar compartment is controlled, then one would predict that pre-assembled flagellar structures, such as radial spokes (Diener et al., 1996) and dyneins arms (Fowkes and Mitchell, 1998), either have flagellar localization signals on one or more of their constituent polypeptides, or that they are escorted through the pore by a carrier – perhaps the IFT particles -- with which they associate. Therefore, the flagellar pore, where non-membrane bound polypeptides dock prior to gaining entrance to the flagellar compartment, would be functionally similar to the nuclear pore.

A flagellar pore complex?

Immunofluorescence microscopy indicates that IFT particle proteins are localized primarily to the base of the Chlamydomonas flagellum, with only a few particles in the flagellum that must represent IFT rafts in transit. Immuno-electron microscopy reveals that the IFT particle proteins are docked at the ends of "transition fibers" that extend from the distal end of the basal body to the plasma membrane at the base of the flagellum. These fibers, which connect each of the nine basal body triplet microtubules to the flagellar membrane, demarcate the boundary between the cytoplasmic and flagellar compartments. The transition fibers thus may be the structural basis for a "flagellar pore complex" (FPC) that limits access to the flagellar compartment (see Figure 3). We propose that flagellar membrane and axonemal proteins synthesized in the cytoplasm are transported to the base of the flagellum, where they are recognized by IFT particle proteins and ushered through the FPC into the flagellar compartment. Transition fibers are present in association with all basal bodies, so it is expected that all cilia and flagella have an FPC that serves as a gateway for admission of specific proteins to the cilium or flagellum.

IFT particle polypeptides in *C. elegans*

Database searching using the Chlamydomonas IFT particle protein sequences revealed a large number of homologues in the nematode C. elegans for which there are known mutations. Interestingly, the mutations are in the Che, Daf, and Osm genes that are required for formation and function of the worm sensory cilia. For example, Osm-1, Osm-5, and Osm-6 encode homologues of the Chlamydomonas IFT172, IFT88 and IFT52 IFT particle proteins (Collet et

al., 1998; Wicks et al., 2000; Qin et al., 2001). Therefore, IFT is essential for the assembly of non-motile sensory cilia in C. elegans. These findings, as well as those on Che-3, which encodes the retrograde IFT motor DHC1b (see above), indicated that the process of IFT has been highly conserved throughout evolution and is likely to be necessary for the assembly of all flagella and cilia, including structures derived from these organelles.

It was presumed that, like the Chlamydomonas IFT particles, the IFT polypeptides initially localized in C. elegans sensory cilia move anterogradely and retrogradely in the cilia. This has now been elegantly demonstrated by fusing sequences encoding GFP to either the kinesin-II non-motor subunit (KAP) gene or to IFT particle polypeptide genes, transforming the constructs into C. elegans, and observing the motility of their products in vivo (Orozco et al., 1999; Signor et al., 1999). The rates of movement of IFT polypeptides and motors were similar to each other, and were likewise similar to those of IFT particles in Chlamydomonas flagella. Recently, several additional C. elegans IFT particle polypeptides, contained in both IFT Complexes A and B, have been tagged with GFP and their motility observed in vivo. Complexes A and B, and the motors that move them, were all shown to translocate at the same rates in neuronal sensory cilia (Qin et al., 2001).

IFT and polycystic kidney disease

Searching the databases with the sequences of the Chlamydomonas IFT particle proteins also revealed close homologues in higher organisms, including mice and humans. The homologue of Chlamydomonas IFT88, termed Tg737 in mouse and man, was of particular interest, because an insertional mutation in this gene causes autosomal recessive polycystic kidney disease (ARPKD) in the mouse (Moyer et al., 1994). In both the mouse and humans, ARPKD involves the formation of numerous cysts in the proximal and collecting tubules of the kidney (Grantham et al., 1996; Moyer et al., 1994). In humans, ARPKD affects up to 1 in 10,000 newborns; most die within a few weeks of birth (Blythe and Ockenden, 1971; Cole et al., 1987). ARPKD also may be responsible for ~1 in 3000 prenatal deaths and still births.

Although the kidney lacks motile cilia, many of the epithelial cells of the kidney collecting ducts and tubules have a single, non-motile '9+0' cilium, called the primary cilium. Primary cilia are in fact present on most cells in the body (see http://www.members.global2000.net/bowser/cilialist.html for a comprehensive list of cells

having primary cilia), but in the kidney they are particularly well developed. The fact that IFT88 is essential for flagellar formation in Chlamydomonas (see above) suggested that its mammalian homologue, Tg737, might be important for formation of the primary cilia in the kidney. Indeed, when kidneys of mice homozygous for the Tg737 insertional mutation were examined by scanning electron microscopy, they were found to be defective in ciliary assembly (Pazour et al., 2000). Whereas wild-type mice had cilia extending several microns into the lumens of the collecting ducts and tubules, the mutant had only short stubs of cilia, just as in the case of the Chlamydomonas IFT88 deletion mutant. These results indicated that the primary cause of ARPKD in the mutant mouse is failure to assemble the primary cilium due to a defect in an IFT particle protein. Subsequent studies have shown that the Tg737 protein is concentrated at the base of the kidney cilia (Taulman et al., 2001), just as IFT particle proteins are located at the base of the flagellum of Chlamydomonas, and that GFP-tagged Tg737 moves anterogradely and retrogradely in the cilia of wild-type kidney cells in culture (Pazour, unpublished results).

The function of the primary cilium is unknown (Alberts et al., 1994). In the kidney, the primary cilium may function as a flowmeter (Schwartz et al., 1997), as an osmometer, or as a chemodetector. Whatever its specific function, the results with the Tg737 mutant mice provide the first evidence that it has a very important role in kidney physiology (Pazour et al., 2000).

A related renal disease, autosomal dominant polycystic kidney disease (ADPKD), which affects up to 1 in 500 individuals, similarly results in the formation of renal cysts, although the symptoms may not be clinically apparent until the patient reaches middle age (Grantham et al., 199_). The primary defects in the most common forms of ADPKD are in genes coding for polycystin 1 (PKD1) and polycystin 2 (PKD2). PKD1 is an integral membrane protein that directly interacts with PKD2 to form a calcium-sensitive cation channel probably acting in a sensory signaling pathway (Emmons and Somlo, 1999; Murcia et al., 1999; Somlo and Ehrlich, 2001). How are the ADPKD polycystins and the defect in kidney cilia assembly in ARPKD related? A clue to this was a report showing that the C. elegans homologues of the vertebrate polycystins are located on the worm's sensory cilia (Barr and Sternberg, 1999), where they also appear to be involved in signal transduction. More recent work indicates that the human polycystins likewise are displayed principally on the ciliary membrane (Pazour et al., manuscript in preparation), suggesting that a defect in one of the polycystins impairs normal functioning of the kidney primary cilium and lead to ADPKD.

IFT reveals a role for primary cilia in polycystic kidney disease

The cells of the proximal and distal collecting tubules of the vertebrate kidney each have a single nonmotile "9 + 0" primary cilium extending from the apical cell surface into the lumen of the tubule. The function of these primary cilia is unknown, but if their assembly is disrupted in the mouse by a defect in a homologue of one of the Chlamydomonas IFT particle polypeptides, autosomal recessive polycystic kidney disease (ARPKD) results and the mice die during gestation or soon after birth (Pazour et al., 2000). Therefore, vertebrate ARPKD occurs when the kidney tubule primary cilia are missing or defective.

In humans the most common genetic deficiencies in a related disorder, autosomal dominant polycystic kidney disease (ADPKD), are in the genes PKD1 and PKD2, which encode polycystin 1 and 2, respectively (Somlo and Ehrlich, 2001). The two polycystins interact to form a calcium-activated cation channel involved in signal transduction (Gonzalez-Perrett et al., 2001).

The reason why either lack of the primary cilia or mutations in the polycystins cause the polycystic kidney phenotype is likely to be because the polycystins on the cell surface are concentrated on the membranes of the kidney primary cilia (Pazour et al., manuscript in preparation). This suggests that a defect in a polycystin impairs functioning of the primary cilium. Therefore, polycystic kidney disease can result either from mutations in the polycystins that are targeted to the primary cilia, or from an inability to form the primary cilia themselves.

IFT and retinal degenerative disease

Primary cilia, or structures derived from primary cilia, also are involved in the development and function of several sensory structures in the vertebrate body, e.g., in the retina, the inner ear, and the nasal epithelium. Inasmuch as IFT probably occurs in all cilia and flagella, it is likely to be important for the assembly and maintenance of these sensory structures/tissues as well. The role of IFT has been most closely examined in retinal rod and cone outer segments. These light sensory dendritic processes, containing the photopigments and light-transducing machinery, initially form from primary cilia (De Robertis, 1956; Tokuyasu and Yamada, 1959); a short '9 + 0' "connecting cilium" remains in the adult as the only path of communication between the outer segment and the inner segment, where protein synthesis occurs (Young, 1976;

Besharse, 1986). Following its formation, the rod outer segment turns over continuously at a high rate; it is estimated that ~2000 opsin molecules per minute are required to maintain the mammalian rod outer segment (Besharse, 1986), and all of this newly synthesized protein is likely to be transported to the outer segment through the connecting cilium. A possible role for IFT in this process was first indicated by the discovery that the IFT motor kinesin-II is present in the connecting cilia of fish rods and cones (Beech et al., 1996). Subsequently, Cre-loxP mutagenesis was used to remove the kinesin-II motor subunit, KIF3A, specifically from photoreceptor cells (Marszalek et al., 2000). In the absence of KIF3A, large quantities of opsin, arrestin and membranes accumulated in the inner segment, and the photoreceptor cells eventually underwent apoptotic cell death. These results implied that kinesin-II, present in the connecting cilium, was powering IFT there and was required for the assembly and continued maintenance of the rod outer segment.

More recently, immunofluorescence microscopy has revealed that several IFT particle proteins are concentrated at the proximal ends, and to a lesser extent the distal ends, of mouse rod connecting cilia (Pazour et al., 2000b), a distribution remarkably similar to that seen for IFT particle proteins in Chlamydomonas. Moreover, in mice homozygous for the insertional mutation in the IFT particle protein Tg737, the rod outer segments develop abnormally, and eventually degenerate, leading to complete disappearance of the rod cells (Pazour et al., manuscript in preparation). These latter results provide very strong independent evidence that IFT occurs in the connecting cilium and has an important role in assembling and maintaining the rod outer segment, presumably by transporting essential proteins from the inner segment to the outer segment (see Figure 5). The degeneration of rod cells resulting from defects in IFT motors and particle proteins is very similar to that observed in retinitis pigmentosa and other human diseases causing progressive blindness (Sung and Tai, 2000; Traboulsi, 1998). Therefore, the genes encoding IFT proteins must now be considered candidate genes for these diseases.

IFT and rod outer segments

Vertebrate rod outer segments are derived from '9 + 0' primary cilia during embryogenesis of the retina. The distal portion of the cilium differentiates into folds that contain the visual pigments and the phototransduction machinery. The proximal portion of the primary cilium remains as the "connecting cilium." The connecting cilium is the only connection

between the rod inner segment, which contains all the cytoplasmic organelles, and the outer segment. In the adult, the outer segment turns over very rapidly due to shedding of membranous disks from its distal tip. Both the initial development of the outer segment and its continued maintenance are dependent on the movement of precursors from their site of synthesis in the inner segment, through the connecting cilium to the outer segment.

Membrane proteins such as rhodopsin that are destined for the outer segment are synthesized in the endoplasmic reticulum, processed through the Golgi apparatus, and then transported in vesicles by cytoplasmic dynein 1 along microtubules that converge at the base of the connecting cilium. There the vesicles fuse with the plasma membrane at a specialized structure termed the periciliary ridge complex (Peters et al., 1983). Immunofluorescence microscopy has revealed a high concentration of IFT particle proteins at the base of the connecting cilium, just as in *Chlamydomonas* (Pazour et al., 2000b). It is hypothesized that the rhodopsin and other outer segment precursors become associated with IFT particles at the base of the connecting cilium, and are transported by kinesin II through the flagellar pore complex and up the cilium to the base of the outer segment, where they are released. The IFT particles are then transported by cytoplasmic dynein 2/1b back down the connecting cilium to the peri-basal body region. Mutations in the carboxyl-terminal portion of rhodopsin cause abnormal transport and/or localization of rhodopsin, resulting in retinal degeneration (Sung and Tai, 2000; Tam et al., 2000). Similarly, mutations in kinesin-II or an IFT particle protein in the mouse result in abnormal localization of rhodopsin and retinal degeneration (Marszalek et al., 2000; Pazour et al., manuscript in preparation). Thus, defects in IFT may be one cause of retinal degeneration in humans.

IFT is required for assembly of nodal cilia

As noted above, deletion of the IFT88 gene in Chlamydomonas completely blocks flagellar assembly. Why, then, are connecting cilia formed at all in rod cells of the mouse homozygous for an insertional mutation in the IFT88 homologue, Tg737? The Tg737 insertional allele used in the above studies results in expression of a smaller-than-normal Tg737 protein, probably due to a defect in mRNA splicing (Taulman et al., 2001; G. Pazour, unpublished result). Apparently, this small product is sufficient to support the assembly of the connecting cilium, but not adequate for full development and long-term maintenance of the outer segment.

Interestingly, complete knockout of the mouse Tg737 gene is embryonic lethal (Murcia et al., 2000); the embryos lack nodal cilia and are defective in left-right axis determination, a hallmark of defects in the nodal cilia (see "Defects in IFT reveal a role for nodal cilia in the development of left-right axis determination"). Therefore, Tg737 in the mouse, like its homologue IFT88 in Chlamydomonas, is essential for cilia formation. As in the case of the connecting cilium, the abnormal Tg737 protein in the original insertional mutant mouse must be sufficient to allow nodal cilia formation and normal embryogenesis. One also can infer that some cilia or structures derived from cilia – viz., kidney primary cilia, rod outer segments – are more sensitive to defects in IFT than other cilia – viz. nodal cilia and respiratory cilia (the latter are relatively unaffected in the original Tg737 insertional mutant mouse.

Defects in IFT reveal a role for embryonic nodal cilia in left-right axis determination

Because IFT is required for the formation of *all* cilia and flagella, animals lacking IFT should be defective in all processes that depend on these organelles. Knockouts of IFT motors and particle proteins in the mouse reveal that one of the earliest roles for cilia in the mammalian embryo is the development of left-right asymmetry. This process appears to depend on the motility of nodal cilia, which resemble primary cilia but are unique in that they exhibit an unusual twirling movement not seen in other primary cilia. In the absence of the IFT motor kinesin-II (Nonaka et al., 1998; Marzalek et al., 1999; Takeda et al., 1999) or the IFT particle protein Tg737 (Murcia et al., 2000), nodal cilia fail to develop. In these embryos, the earliest molecular markers of left-right asymmetry, which normally are expressed only in the left lateral plate mesoderm shortly after development of the nodal cilia, are expressed bilaterally (Nonaka et al., 1998; Marszalek et al., 1999). Subsequently, the embryos undergo random laterality of heart looping, so that in half of the embryos the heart is on the wrong side of the midline, a condition known as *situs inversus* and frequently observed in humans with defects in cilia function (Afzelius and Mossberg, 1995). It has been proposed that the twirling movement of the nodal cilia sets up a gradient of a morphogen in the extraembryonic fluid across the node, resulting in left-right axis determination (Okada et al., 1999; Supp et al., 1999). In support of such a role, embryos of mice with a defect in left-right dynein, which may be the motor powering nodal cilia movement, have non-motile nodal cilia and situs inversus (Okada et al., 1999; Supp et al., 1999). (Modified from Supp et al., 2000)

Further Examples of IFT Function

Intraflagellar Transport (IFT) is essential for both the initial assembly and subsequent maintenance of all motile (9 + 2) and sensory (9 + 0) cilia/flagella. The research shows that IFT is required for the movement of structural proteins, which form the microtubular axoneme, to the flagellar tip, where assembly occurs. For those 9 + 2 cilia and flagella where motility is the principal function, e.g. flagella of sperm, and cilia of the respiratory tract, oviduct, fimbriae, efferent ducts of the testis, and ependymal cells lining the cavities of the spinal cord and brain, the motility function is dependent on proper assembly and maintenance of the microtubule-containing axoneme. When the nodal cilia cannot form in the embryo, developmental diseases such as situs inversus result. IFT also is important for movement and maintenance of proteins in the flagellar and ciliary membranes. For example, the membrane-associated proteins involved in phototransduction in the vertebrate retina are in the rod and cone outer segments, which are modified cilia. Finally, IFT is important in the movement of signals from the cilium to the cell body, as in the case of 9 + 0 primary cilia and sensory cilia which monitor the environment and relay that information to the cell or nervous system.

II. DNA and Polypeptide Sequences

The present invention is based, at least in part, on the discovery or characterization of a variety of cDNA and polypeptide molecules, and sequences inferred from or homologous to them, which encode proteins which are herein designated IFT 20, 27, 46, 52, 57, 72, 88, 122, and139, and Che-2. Sequences from Chlamydomonas , Caenorhabditis elegans, mouse, and human are included in the invention. These proteins exhibit a variety of physiological activities, and are included in a single application for the sake of convenience. It is understood that the allowability or non-allowability of claims directed to one of these proteins has no bearing on the allowability of claims directed to the others. The characteristics of each of these proteins and the cDNAs encoding them are described separately in the ensuing sections. In addition to the full length mature and immature human proteins described in the following sections, the invention includes fragments, derivatives, and variants of these proteins, as described herein. These proteins, fragments, derivatives, and variants are collectively referred to herein as polypeptides of the invention or proteins of the invention.

IFT27 is a member of the Rab family of small GTPases. Rab GTPases and their effectors are generally involved in membrane vesicle formation, vesicle and organelle motility and transport, and tethering of vesicles to their target compartment (Zerial and McBride, *Nature Reviews Molecular Cell Biology*, 2:107-119, 2001). Known Rabs cycle between an "active" state in which GTP is bound, and an "inactive" state in which GDP is bound. Thus they serve as molecular "switches."

An "isolated nucleic acid molecule" is a nucleic acid molecule that is separated from the 5' and 3' coding sequences with which it is immediately contiguous in the naturally occurring genome of an organism. Isolated nucleic acid molecules include nucleic acid molecules that are not naturally occurring, e.g., nucleic acid molecules created by recombinant DNA techniques. Nucleic acid molecules include both RNA and DNA, including cDNA, genomic DNA, and synthetic (e.g., chemically synthesized) DNA. Where single-stranded, the nucleic acid molecule may be a sense strand or an antisense strand.

As used herein, a "signal sequence" includes a peptide of at least about 15 or 20 amino acid residues in length which occurs at the N-terminus of secretory and membrane-bound proteins and which contains at least about 70% hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a signal sequence contains at least about 10 to 40 amino acid residues, preferably about 19-34 amino acid residues, and has at least about 60-80%, more preferably at least about 65-75%, and more preferably at least about 70% hydrophobic residues. A signal sequence serves to direct a protein containing such a sequence to a lipid bilayer. A signal sequence is usually cleaved during processing of the mature protein.

The term "purified" as used herein refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized.

Polypeptides or other compounds of interest are said to be "substantially pure" when they are within preparations that are at least 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity can be measured by any appropriate standard method, for example, by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

Where a particular polypeptide or nucleic acid molecule is said to have a specific percent identity to a reference polypeptide or nucleic acid molecule of a defined length, the percent identity is relative to the reference polypeptide or nucleic acid molecule. Thus, a peptide that is 50% identical to a reference polypeptide that is 100 amino acids long can be a 50 amino acid polypeptide that is completely identical to a 50 amino acid long portion of the reference polypeptide. It might also be a 100 amino acid long polypeptide that is 50% identical to the reference polypeptide over its entire length. Of course, many other polypeptides will meet the same criteria. The same rule applies for nucleic acid molecules.

For polypeptides, the length of the reference polypeptide sequence will generally be at least 16 amino acids, preferably at least 20 amino acids, more preferably at least 25 amino acids, and most preferably 35 amino acids, 50 amino acids, or 100 amino acids. For nucleic acids, the length of the reference nucleic acid sequence will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides or 300 nucleotides.

In the case of polypeptide sequences which are less than 100% identical to a reference sequence, the non-identical positions are preferably, but not necessarily, conservative substitutions for the reference sequence. Conservative substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine, and leucine; aspartic acid and glutamic acid; asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions (e.g., overlapping positions) x 100). Preferably, the two sequences are the same length.

The determination of percent homology between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to IFT 20, 27, 46, 52, 57, 72, 88, 122, or139, or Che-2 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to IFT 20, 27, 46, 52, 57, 72, 88, 122, or139, or Che-2 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search that detects distant relationships between molecules. Id. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. *See* http://www.ncbi.nlm.nih.gov. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the local homology algorithm of Smith and Waterman (Advances in Applied Mathematics 2: 482-489 (1981)). Such an algorithm is incorporated into the BestFit program, which is part of the Wisconsin™ package, and is used to find the best segment of similarity between two sequences. BestFit reads a scoring matrix that contains values for every possible GCG symbol match. The program uses these values to construct a path matrix that represents the entire surface of comparison with a score at every position for the best possible alignment to that point. The quality score for the best alignment to any point is equal to the sum of the scoring matrix values of the matches in that alignment, less the gap creation penalty multiplied by the number of gaps in that alignment, less the gap extension penalty multiplied by the total length of all gaps in that

alignment. The gap creation and gap extension penalties are set by the user. If the best path to any point has a negative value, a zero is put in that position.

After the path matrix is complete, the highest value on the surface of comparison represents the end of the best region of similarity between the sequences. The best path from this highest value backwards to the point where the values revert to zero is the alignment shown by BestFit. This alignment is the best segment of similarity between the two sequences. Further documentation can be found at http://ir.ucdavis.edu/GCGhelp/bestfit.html#algorithm.

Additional algorithms for sequence analysis are known in the art and include ADVANCE and ADAM as described in Torellis and Robotti (1994) Comput. Appl. Biosci., 10:3-5; and FASTA described in Pearson and Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-8. Within FASTA, ktup is a control option that sets the sensitivity and speed of the search. If ktup=2, similar regions in the two sequences being compared are found by looking at pairs of aligned residues; if ktup=1, single aligned amino acids are examined. ktup can be set to 2 or 1 for protein sequences, or from 1 to 6 for DNA sequences. The default if ktup is not specified is 2 for proteins and 6 for DNA. For a further description of FASTA parameters, see http://bioweb.pasteur.fr/docs/man/man/fasta.1.html#sect2, the contents of which are incorporated herein by reference.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically exact matches are counted.

As used herein, the phrase "allelic variant" refers to a nucleotide sequence that occurs at a given locus or to a polypeptide encoded by the nucleotide sequence. Allelic variants of any of these genes can be identified by sequencing the corresponding chromosomal portion at the indication location in multiple individuals.

Nucleic Acids Encoding IFT Particle Proteins

The invention encompasses nucleic acids that have a sequence that is substantially identical to the nucleic acid sequence of Chlamydomonas IFT particle protein genes 20, 27, 46, 52, 57, 72, 88, 122, or139, or Che-2, as well as homologous mouse and human sequences. A nucleic acid sequence which is substantially identical to a given reference nucleic acid sequence is hereby defined as a nucleic acid having a sequence that has at least 85%, preferably 90%, and more preferably 95%, 98%, 99% or more identity to the sequence of the given reference nucleic acid sequence.

The IFT 20, 27, 46, 52, 57, 72, 88, 122, or 139, or Che-2 nucleic acid molecules of the invention can be cDNA, genomic DNA, synthetic DNA, or RNA, and can be double-stranded or single-stranded (*i.e.*, either a sense or an antisense strand). Fragments of these molecules are also considered within the scope of the invention, and can be produced, for example, by the polymerase chain reaction (PCR) or generated by treatment with one or more restriction endonucleases. A ribonucleic acid (RNA) molecule can be produced by *in vitro* transcription.

The nucleic acid molecules of the invention can contain naturally occurring sequences, or sequences that differ from those that occur naturally, but, due to the degeneracy of the genetic code, encode the same polypeptide. In addition, these nucleic acid molecules are not limited to sequences that only encode polypeptides, and thus, can include some or all of the non-coding sequences that lie upstream or downstream from a coding sequence.

The nucleic acid molecules of the invention can be synthesized (for example, by phosphoramidite-based synthesis) or obtained from a biological cell, such as the cell of a mammal. Thus, the nucleic acids can be those of a human, mouse, rat, guinea pig, cow, sheep, horse, pig, rabbit, monkey, dog, or cat. Combinations or modifications of the nucleotides within these types of nucleic acids are also encompassed.

In addition, the isolated nucleic acid molecules of the invention encompass fragments that are not found as such in the natural state. Thus, the invention encompasses recombinant molecules, such as those in which a nucleic acid molecule (for example, an isolated nucleic acid molecule encoding IFT 20, 27, 46, 52, 57, 72, 88, 122, or139, or Che-2) is incorporated into a vector (for example, a plasmid or viral vector) or into the genome of a heterologous cell (or the genome of a homologous cell, at a position other than the natural chromosomal location). Recombinant nucleic acid molecules and uses therefor are discussed further below.

In the event the nucleic acid molecules of the invention encode or act as antisense molecules, they can be used for example, to regulate translation of mRNA of the invention. Techniques associated with detection or regulation of expression of nucleic acids or polypeptides of the invention are well known to skilled artisans and can be used to diagnose and/or treat disorders associated with aberrant expression of nucleic acids or polypeptides of the invention.

The invention also encompasses nucleic acid molecules that hybridize under stringent conditions to a nucleic acid molecule encoding a polypeptide of the invention. The cDNA sequences described herein can be used to identify these hybridizing nucleic acids, which include, for example,

nucleic acids that encode homologous polypeptides in other species, and splice variants of the genes of the invention in humans or other mammals. Accordingly, the invention features methods of detecting and isolating these nucleic acid molecules. Using these methods, a sample (for example, a nucleic acid library, such as a cDNA or genomic library) is contacted (or "screened") with a probe specific to a nucleotide of the invention. The probe will selectively hybridize to nucleic acids encoding related polypeptides (or to complementary sequences thereof). The probe, which can contain at least 25 (for example, 25, 50, 100, or 200 nucleotides) can be produced using any of several standard methods (see, for example, Ausubel et al.,"Current Protocols in Molecular Biology, Vol. I," Green Publishing Associates, Inc., and John Wiley & Sons, Inc., NY, 1989). For example, the probe can be generated using PCR amplification methods in which oligonucleotide primers are used to amplify a nucleic acid sequence specific to a nucleic acid of the invention that can be used as a probe to screen a nucleic acid library and thereby detect nucleic acid molecules (within the library) that hybridize to the probe.

One single-stranded nucleic acid is said to hybridize to another if a duplex forms between them. This occurs when one nucleic acid contains a sequence that is the reverse and complement of the other (this same arrangement gives rise to the natural interaction between the sense and antisense strands of DNA in the genome and underlies the configuration of the "double helix"). Complete complementarity between the hybridizing regions is not required in order for a duplex to form; it is only necessary that the number of paired bases is sufficient to maintain the duplex under the hybridization conditions used.

Typically, hybridization conditions are of low to moderate stringency. These conditions favor specific interactions between completely complementary sequences, but allow some non-specific interaction between less than perfectly matched sequences to occur as well. After hybridization, the nucleic acids can be "washed" under moderate or high conditions of stringency to dissociate duplexes that are bound together by some non-specific interaction (the nucleic acids that form these duplexes are thus not completely complementary).

As is known in the art, the optimal conditions for washing are determined empirically, often by gradually increasing the stringency. The parameters that can be changed to affect stringency include, primarily, temperature and salt concentration. In general, the lower the salt concentration and the higher the temperature, the higher the stringency. Washing can be initiated at a low temperature (for example, room temperature) using a solution containing a salt concentration that is

equivalent to or lower than that of the hybridization solution. Subsequent washing can be carried out using progressively warmer solutions having the same salt concentration. As alternatives, the salt concentration can be lowered and the temperature maintained in the washing step, or the salt concentration can be lowered and the temperature increased. Additional parameters can also be altered. For example, use of a destabilizing agent, such as formamide, alters the stringency conditions.

In reactions where nucleic acids are hybridized, the conditions used to achieve a given level of stringency will vary. There is not one set of conditions, for example, that will allow duplexes to form between all nucleic acids that are 85% identical to one another; hybridization also depends on unique features of each nucleic acid. The length of the sequence, the composition of the sequence (for example, the content of purine-like nucleotides versus the content of pyrimidine-like nucleotides) and the type of nucleic acid (for example, DNA or RNA) affect hybridization. An additional consideration is whether one of the nucleic acids is immobilized (for example, on a filter).

An example of a progression from lower to higher stringency conditions is the following, where the salt content is given as the relative abundance of SSC (a salt solution containing sodium chloride and sodium citrate; 2X SSC is 10-fold more concentrated than 0.2X SSC). Nucleic acids are hybridized at 42°C in 2X SSC/0.1% SDS (sodium dodecylsulfate; a detergent) and then washed in 0.2X SSC/0.1% SDS at room temperature (for conditions of low stringency); 0.2X SSC/0.1% SDS at 42°C (for conditions of moderate stringency); and 0.1X SSC at 68°C (for conditions of high stringency). Washing can be carried out using only one of the conditions given, or each of the conditions can be used (for example, washing for 10-15 minutes each in the order listed above). Any or all of the washes can be repeated. As mentioned above, optimal conditions will vary and can be determined empirically.

A second set of conditions that are considered "stringent conditions" are those in which hybridization is carried out at 50°C in Church buffer (7% SDS, 0.5% $NaHPO_4$, 1 M EDTA, 1% BSA) and washing is carried out at 50°C in 2X SSC.

Once detected, the nucleic acid molecules can be isolated by any of a number of standard techniques (see, for example, Sambrook et al., "Molecular Cloning, A Laboratory Manual," 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

The invention also encompasses: (a) expression vectors that contain any of the foregoing coding sequences (related to a polypeptide of the invention) and/or their complements (that is,

"antisense" sequence); (b) expression vectors that contain any of the foregoing coding sequences (related to a polypeptide of the invention) operatively associated with a regulatory element (examples of which are given below) that directs the expression of the coding sequences; (c) expression vectors containing, in addition to sequences encoding a polypeptide of the invention, nucleic acid sequences that are unrelated to nucleic acid sequences encoding a polypeptide of the invention, such as molecules encoding a reporter or marker; and (d) genetically engineered host cells that contain any of the foregoing expression vectors and thereby express the nucleic acid molecules of the invention in the host cell.

Recombinant nucleic acid molecules can contain a sequence encoding a soluble polypeptide of the invention; mature polypeptide of the invention; or polypeptide of the invention having an added or endogenous signal sequence. A full length polypeptide of the invention; a domain of a polypeptide of the invention; or a fragment thereof may be fused to additional polypeptides, as described below. Similarly, the nucleic acid molecules of the invention can encode the mature form of a polypeptide of the invention or a form that encodes a polypeptide that facilitates secretion. In the latter instance, the polypeptide is typically referred to as a proprotein, which can be converted into an active form by removal of the signal sequence, for example, within the host cell. Proproteins can be converted into the active form of the protein by removal of the inactivating sequence.

The regulatory elements referred to above include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements, which are known to those skilled in the art, and which drive or otherwise regulate gene expression. Such regulatory elements include but are not limited to the cytomegalovirus hCMV immediate early gene, the early or late promoters of SV40 adenovirus, the lac system, the trp system, the TAC system, the TRC system, the major operator and promoter regions of phage A, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase, the promoters of acid phosphatase, and the promoters of the yeast α-mating factors.

Similarly, the nucleic acid can form part of a hybrid gene encoding additional polypeptide sequences, for example, sequences that function as a marker or reporter. Examples of marker or reporter genes include β-lactamase, chloramphenicol acetyltransferase (CAT), adenosine deaminase (ADA), aminoglycoside phosphotransferase (neo^r, $G418^r$), dihydrofolate reductase (DHFR), hygromycin-B-phosphotransferase (HPH), thymidine kinase (TK), lacZ (encoding β-galactosidase), and xanthine guanine phosphoribosyltransferase (XGPRT). As with many of the standard

procedures associated with the practice of the invention, skilled artisans will be aware of additional useful reagents, for example, of additional sequences that can serve the function of a marker or reporter. Generally, the hybrid polypeptide will include a first portion and a second portion; the first portion being a polypeptide of the invention and the second portion being, for example, the reporter described above or an immunoglobulin constant region.

The expression systems that may be used for purposes of the invention include, but are not limited to, microorganisms such as bacteria (for example, *E. coli* and *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA, or cosmid DNA expression vectors containing the nucleic acid molecules of the invention; yeast (for example, *Saccharomyces* and *Pichia*) transformed with recombinant yeast expression vectors containing the nucleic acid molecules of the invention (preferably containing the nucleic acid sequence encoding a polypeptide of the invention); insect cell systems infected with recombinant virus expression vectors (for example, baculovirus) containing the nucleic acid molecules of the invention; plant cell systems infected with recombinant virus expression vectors (for example, cauliflower mosaic virus (CaMV) and tobacco mosaic virus (TMV)) or transformed with recombinant plasmid expression vectors (for example, Ti plasmid) containing nucleotide sequences of nucleic acids of the invention; or mammalian cell systems (for example, COS, CHO, BHK, 293, VERO, HeLa, MDCK, WI38, and NIH 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (for example, the metallothionein promoter) or from mammalian viruses (for example, the adenovirus late promoter and the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions containing polypeptides of the invention or for raising antibodies to those polypeptides, vectors that are capable of directing the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited to, the *E. coli* expression vector pUR278 (Ruther et al., *EMBO J.* 2:1791, 1983), in which the coding sequence of the insert may be ligated individually into the vector in frame with the lacZ coding region so that a fusion protein is produced; pIN vectors (Inouye and Inouye, *Nucleic Acids Res.* 13:3101-3109, 1985; Van Heeke and Schuster, *J. Biol. Chem.* 264:5503-5509, 1989); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general,

such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhidrosis virus (AcNPV) can be used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The coding sequence of the insert may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed. (for example, see Smith et al., *J. Virol.* 46:584, 1983; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the nucleic acid molecule of the invention may be ligated to an adenovirus transcription/translation control complex, for example, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (for example, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing a gene product of the invention in infected hosts (for example, see Logan and Shenk, *Proc. Natl. Acad. Sci. USA* 81:3655-3659, 1984). Specific initiation signals may also be required for efficient translation of inserted nucleic acid molecules. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate

transcription enhancer elements, transcription terminators, etc. (see Bittner et al., *Methods in Enzymol.* 153:516-544, 1987).

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (for example, glycosylation) and processing (for example, cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. The mammalian cell types listed above are among those that can serve as suitable host cells.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express the sequences of nucleic acids or polypeptides of the invention described above may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (for example, promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci that in turn can be cloned and expanded into cell lines. This method can advantageously be used to engineer cell lines that express nucleic acids or polypeptides of the invention. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the gene product.

A number of selection systems can be used. For example, the herpes simplex virus thymidine kinase (Wigler, et al., *Cell* 11:223, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, *Proc. Natl. Acad. Sci. USA* 48:2026, 1962), and adenine phosphoribosyltransferase (Lowy, et al., *Cell* 22:817, 1980) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells, respectively. Also, anti-metabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., *Proc.*

Natl. Acad. Sci. USA 77:3567, 1980; O'Hare et al., *Proc. Natl. Acad. Sci. USA* 78:1527, 1981); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072, 1981); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin et al., *J. Mol. Biol.* 150:1, 1981); and hygro, which confers resistance to hygromycin (Santerre et al., *Gene* 30:147, 1984).

The nucleic acid molecules of the invention are useful for diagnosis of disorders associated with aberrant expression of nucleic acid molecules of the invention are also useful in genetic mapping and chromosome identification.

IFT Particle Polypeptides

The invention also includes polypeptides that have a sequence that is substantially identical to the amino acid sequence of Chlamydomonas IFT particle polypeptides 20, 27, 46, 52, 57, 72, 88, 122, or 139, or Che-2. A polypeptide which is "substantially identical" to a given reference polypeptide is a polypeptide having a sequence that has at least 85%, preferably 90%, and more preferably 95%, 98%, 99% or more identity to the sequence of the given reference polypeptide sequence.

The terms "protein" and "polypeptide" are used herein interchangably to describe any chain of amino acids, regardless of length or post-translational modification (for example, glycosylation or phosphorylation). Thus, the term "IFT 20, 27, 46, 52, 57, 72, 88, 122, or 139, or Che-2 polypeptide" includes: full-length, naturally occurring protein of the invention; recombinantly or synthetically produced polypeptide that corresponds to a full-length naturally occurring protein of the invention; or particular domains or portions of the naturally occurring protein. The term also encompasses mature a polypeptide of the invention that has an added amino-terminal methionine (useful for expression in prokaryotic cells).

The polypeptides of the invention described herein are those encoded by any of the nucleic acid molecules described above and include fragments, mutants, truncated forms, and fusion proteins of polypeptides of the invention. These polypeptides can be prepared for a variety of uses, including but not limited to the generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products or compounds that can modulate the activity or expression of nucleic acids or polypeptides of the invention, and as pharmaceutical reagents useful for the treatment of

disorders associated with aberrant expression or activity of nucleic acids or polypeptides of the invention.

Preferred polypeptides are substantially pure polypeptides of the invention, including those that correspond to the polypeptide with an intact signal sequence, and the secreted form of the polypeptide. Especially preferred are polypeptides that are soluble under normal physiological conditions.

The invention also encompasses polypeptides that are functionally equivalent to polypeptides of the invention. These polypeptides are equivalent to polypeptides of the invention in that they are capable of carrying out one or more of the functions of polypeptides of the invention in a biological system. Preferred polypeptides of the invention have 20%, 40%, 50%, 75%, 80%, or even 90% of one or more of the biological activities of the full-length, mature human form of polypeptides of the invention. Such comparisons are generally based on an assay of biological activity in which equal concentrations of the polypeptides are used and compared. The comparison can also be based on the amount of the polypeptide required to reach 50% of the maximal stimulation obtainable.

Functionally equivalent proteins can be those, for example, that contain additional or substituted amino acid residues. Substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. Conservative substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine, and leucine; aspartic acid and glutamic acid; asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine.

Polypeptides that are functionally equivalent to polypeptides of the invention can be made using random mutagenesis techniques well known to those skilled in the art. It is more likely, however, that such polypeptides will be generated by site-directed mutagenesis (again using techniques well known to those skilled in the art). These polypeptides may have increased functionality or decreased functionality.

To design functionally equivalent polypeptides, it is useful to distinguish between conserved positions and variable positions. This can be done by aligning the amino acid sequence of a protein of the invention from one species with its homolog from another species. Skilled artisans will recognize that conserved amino acid residues are more likely to be necessary for preservation of function. Thus, it is preferable that conserved residues are not altered.

Mutations within the coding sequence of nucleic acid molecules of the invention can be made to generate variant genes that are better suited for expression in a selected host cell. For example, N-linked glycosylation sites can be altered or eliminated to achieve, for example, expression of a homogeneous product that is more easily recovered and purified from yeast hosts which are known to hyperglycosylate N-linked sites. To this end, a variety of amino acid substitutions at one or both of the first or third amino acid positions of any one or more of the glycosylation recognition sequences which occur, and/or an amino acid deletion at the second position of any one or more of such recognition sequences, will prevent glycosylation at the modified tripeptide sequence (see, for example, Miyajima et al., *EMBO J.* 5:1193, 1986).

The polypeptides of the invention can be expressed fused to another polypeptide, for example, a marker polypeptide or fusion partner. For example, the polypeptide can be fused to a hexa-histidine tag to facilitate purification of bacterially expressed protein or a hemagglutinin tag to facilitate purification of protein expressed in eukaryotic cells.

A fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (*Proc. Natl. Acad. Sci. USA* 88: 8972-8976, 1991). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+}·nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

The polypeptides of the invention can be chemically synthesized (for example, see Creighton, "Proteins: Structures and Molecular Principles," W.H. Freeman & Co., NY, 1983), or, perhaps more advantageously, produced by recombinant DNA technology as described herein. For additional guidance, skilled artisans may consult Ausubel et al. (*supra*), Sambrook et al. ("Molecular Cloning, A Laboratory Manual," Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989), and, particularly for examples of chemical synthesis Gait, M.J. Ed. ("Oligonucleotide Synthesis," IRL Press, Oxford, 1984).

The invention also features polypeptides that interact with nucleic acids or polypeptides of the invention (and the genes that encode them) and thereby alter the function of nucleic acids or polypeptides of the invention. Interacting polypeptides can be identified using methods known to

those skilled in the art. One suitable method is the "two-hybrid system," which detects protein interactions *in vivo* (Chien et al., *Proc. Natl. Acad. Sci. USA*, 88:9578, 1991). A kit for practicing this method is available from Clontech (Palo Alto, CA).

Transgenic animals

Polypeptides of the invention can also be expressed in transgenic animals. These animals represent a model system for the study of disorders that are caused by or exacerbated by overexpression or underexpression of nucleic acids or polypeptides of the invention, and for the development of therapeutic agents that modulate the expression or activity of nucleic acids or polypeptides of the invention.

Transgenic animals can be farm animals (pigs, goats, sheep, cows, horses, rabbits, and the like), rodents (such as rats, guinea pigs, and mice), non-human primates (for example, baboons, monkeys, and chimpanzees), and domestic animals (for example, dogs and cats). Transgenic mice are especially preferred.

Any technique known in the art can be used to introduce an IFT 20, 27, 46, 52, 57, 72, 88, 122, or139, or Che-2 transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten et al., *Proc. Natl. Acad. Sci., USA* 82:6148, 1985); gene targeting into embryonic stem cells (Thompson et al., *Cell* 56:313, 1989); and electroporation of embryos (Lo, *Mol. Cell. Biol.* 3:1803, 1983).

The present invention provides for transgenic animals that carry a transgene of the invention in all their cells, as well as animals that carry a transgene in some, but not all of their cells. That is, the invention provides for mosaic animals. The transgene can be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene can also be selectively introduced into and activated in a particular cell type (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89:6232, 1992). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

When it is desired that the transgene of the invention be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be used, vectors containing some nucleotide sequences homologous to an endogenous gene of the invention are designed for the purpose of integrating, via homologous recombination with chromosomal

sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene also can be selectively introduced into a particular cell type, thus inactivating the endogenous gene of the invention in only that cell type (Gu et al., *Science* 265:103, 1984). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. These techniques are useful for preparing "knock outs" lacking a functional gene.

Once transgenic animals have been generated, the expression of the recombinant gene of the invention can be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to determine whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR. Biological samples can also be evaluated immunocytochemically using antibodies specific for the product of the transgene of the invention. Samples of tissue expressing the gene of the invention can also be evaluated immunocytochemically using antibodies specific for the product of the transgene of the invention.

For a review of techniques that can be used to generate and assess transgenic animals, skilled artisans can consult Gordon (*Intl. Rev. Cytol.* 115:171-229, 1989), and may obtain additional guidance from, for example: Hogan et al. "Manipulating the Mouse Embryo" (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1986; Krimpenfort et al., *Bio/Technology* 9:86, 1991; Palmiter et al., *Cell* 41:343, 1985; Kraemer et al., "Genetic Manipulation of the Early Mammalian Embryo," Cold Spring Harbor Press, Cold Spring Harbor, NY, 1985; Hammer et al., *Nature* 315:680, 1985; Purcel et al., *Science*, 244:1281, 1986; Wagner et al., U.S. Patent No. 5,175,385; and Krimpenfort et al., U.S. Patent No. 5,175,384 (the latter two publications are hereby incorporated by reference).

Anti-IFT 20, 27, 46, 52, 57, 72, 88, 122, or 139, or Che-2 Antibodies

Human polypeptides of the invention (or immunogenic fragments or analogs) can be used to raise antibodies useful in the invention; such polypeptides can be produced by recombinant techniques or synthesized (see, for example, "Solid Phase Peptide Synthesis," *supra*; Ausubel *et al.*, *supra*). In general, the peptides can be coupled to a carrier protein, such as KLH, as described in

Ausubel et al., *supra*, mixed with an adjuvant, and injected into a host mammal. Antibodies can be purified by peptide antigen affinity chromatography.

In particular, various host animals can be immunized by injection with a polypeptide of the invention. Host animals include rabbits, mice, guinea pigs, and rats. Various adjuvants that can be used to increase the immunological response depend on the host species and include Freund's adjuvant (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Potentially useful human adjuvants include BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Polyclonal antibodies are heterogeneous populations of antibody molecules that are contained in the sera of the immunized animals.

Antibodies within the invention therefore include polyclonal antibodies and, in addition, monoclonal antibodies, humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, and molecules produced using a Fab expression library.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, can be prepared using the polypeptides of the invention described above and standard hybridoma technology (see, for example, Kohler et al., *Nature* 256:495, 1975; Kohler et al., *Eur. J. Immunol.* 6:511, 1976; Kohler et al., *Eur. J. Immunol.* 6:292, 1976; Hammerling et al., "Monoclonal Antibodies and T Cell Hybridomas," Elsevier, NY, 1981; Ausubel et al., *supra*).

In particular, monoclonal antibodies can be obtained by any technique that provides for the production of antibody molecules by continuous cell lines in culture such as described in Kohler et al., *Nature* 256:495, 1975, and U.S. Patent No. 4,376,110; the human B-cell hybridoma technique (Kosbor et al., *Immunology Today* 4:72, 1983; Cole et al., *Proc. Natl. Acad. Sci. USA* 80:2026, 1983), and the EBV-hybridoma technique (Cole et al., "Monoclonal Antibodies and Cancer Therapy," Alan R. Liss, Inc., pp. 77-96, 1983). Such antibodies can be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated *in vitro* or *in vivo*. The ability to produce high titers of mAbs *in vivo* makes this a particularly useful method of production.

Once produced, polyclonal or monoclonal antibodies are tested for specific recognition of polypeptides of the invention by Western blot or immunoprecipitation analysis by standard methods, e.g., as described in Ausubel et al., *supra*. Antibodies that specifically recognize and bind to polypeptides of the invention are useful in the invention. For example, such antibodies can be used

in an immunoassay to monitor the level of a polypeptide of the invention produced by a mammal (for example, to determine the amount or subcellular location of a polypeptide of the invention).

Preferably, antibodies of the invention are produced using fragments of the protein of the invention that lie outside highly conserved regions and appear likely to be antigenic, by criteria such as high frequency of charged residues. In one specific example, such fragments are generated by standard techniques of PCR, and are then cloned into the pGEX expression vector (Ausubel et al., *supra*). Fusion proteins are expressed in *E. coli* and purified using a glutathione agarose affinity matrix as described in Ausubel, *et al.*, *supra*.

In some cases it may be desirable to minimize the potential problems of low affinity or specificity of antisera. In such circumstances, two or three fusions can be generated for each protein, and each fusion can be injected into at least two rabbits. Antisera can be raised by injections in a series, preferably including at least three booster injections.

Antisera may also checked for its ability to immunoprecipitate recombinant proteins of the invention or control proteins, such as glucocorticoid receptor, CAT, or luciferase.

The antibodies can be used, for example, in the detection of the polypeptide of the invention in a biological sample as part of a diagnostic assay. Antibodies also can be used in a screening assay to measure the effect of a candidate compound on expression or localization of a polypeptide of the invention. Additionally, such antibodies can be used in conjunction with the gene therapy techniques described to, for example, evaluate normal and/or genetically engineered cells that express nucleic acids or polypeptides of the invention prior to their introduction into the patient. Such antibodies additionally can be used in a method for inhibiting abnormal activity of nucleic acids or polypeptides of the invention.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., *Proc. Natl. Acad. Sci. USA*, 81:6851, 1984; Neuberger et al., *Nature*, 312:604, 1984; Takeda et al., *Nature*, 314:452, 1984) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent

administration are often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank et al. ((1997) J. Acquired Immune Deficiency Syndromes and Human Retrovirology 14:193).

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent Nos. 4,946,778, 4,946,778, and 4,704,692) can be adapted to produce single chain antibodies against polypeptides of the invention. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments that recognize and bind to specific epitopes can be generated by known techniques. For example, such fragments include but are not limited to $F(ab')_2$ fragments that can be produced by pepsin digestion of the antibody molecule, and Fab fragments that can be generated by reducing the disulfide bridges of $F(ab')_2$ fragments. Alternatively, Fab expression libraries can be constructed (Huse et al., *Science*, 246:1275, 1989) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Antibodies to polypeptides of the invention can, in turn, be used to generate anti-idiotype antibodies that resemble a portion of the protein of the invention using techniques well known to those skilled in the art (see, e.g., Greenspan et al., *FASEB J.* 7:437, 1993; Nissinoff, *J. Immunol.* 147:2429, 1991). For example, antibodies that bind to the protein of the invention and competitively inhibit the binding of a binding partner of the protein can be used to generate anti-idiotypes that resemble a binding partner binding domain of the protein and, therefore, bind and neutralize a binding partner of the protein. Such neutralizing anti-idiotypic antibodies or Fab fragments of such anti-idiotypic antibodies can be used in therapeutic regimens.

Antibodies can be humanized by methods known in the art. For example, monoclonal antibodies with a desired binding specificity can be commercially humanized (Scotgene, Scotland; Oxford Molecular, Palo Alto, CA). Fully human antibodies, such as those expressed in transgenic animals are also features of the invention (Green et al., *Nature Genetics* 7:13-21, 1994; see also U.S. Patents 5,545,806 and 5,569,825, both of which are hereby incorporated by reference).

The methods described herein in which anti-polypeptide-of-the-invention antibodies are employed may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific polypeptide-of-the-invention antibody reagent described herein, which may be

conveniently used, for example, in clinical settings, to diagnose patients exhibiting symptoms of disorders associated with aberrant expression of nucleic acids or polypeptides of the invention.

An antibody (or fragment thereof) can be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent, or a radioactive agent (e.g., a radioactive metal ion). Cytotoxins and cytotoxic agents include any agent that is detrimental to cells. Examples of such agents include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, and 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin {formerly designated daunomycin} and doxorubicin), antibiotics (e.g., dactinomycin {formerly designated actinomycin}, bleomycin, mithramycin, and anthramycin), and anti-mitotic agents (e.g., vincristine and vinblastine).

Conjugated antibodies of the invention can be used for modifying a given biological response, the drug moiety not being limited to classical chemical therapeutic agents. For example, the drug moiety can be a protein or polypeptide possessing a desired biological activity. Such proteins include, for example, toxins such as abrin, ricin A, Pseudomonas exotoxin, or diphtheria toxin; proteins such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; and biological response modifiers such as lymphokines, interleukin-1, interleukin-2, interleukin-6, granulocyte macrophage colony stimulating factor, granulocyte colony stimulating factor, or other growth factors.

Techniques for conjugating a therapeutic moiety to an antibody are well known (see, e.g., Arnon et al., 1985, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al., Eds., Alan R. Liss, Inc. pp. 243-256; Hellstrom et al., 1987, "Antibodies For Drug Delivery", in Controlled Drug Delivery, 2nd ed., Robinson et al., Eds., Marcel Dekker, Inc., pp. 623-653; Thorpe, 1985, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al., Eds., pp. 475-506; "Analysis, Results, And Future Prospective

Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al., Eds., Academic Press, pp. 303-316, 1985; and Thorpe et al., 1982, Immunol. Rev., 62:119-158). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

Antisense Nucleic Acids

Treatment regimes based on an "antisense" approach involve the design of oligonucleotides (either DNA or RNA) that are complementary to mRNA of the invention. These oligonucleotides bind to the complementary mRNA transcripts of the invention and prevent translation. Absolute complementarity, although preferred, is not required. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarily to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarily and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs recently have been shown to be effective at inhibiting translation of mRNAs as well (Wagner, *Nature* 372:333, 1984). Thus, oligonucleotides complementary to either the 5' or 3' non-translated, non-coding regions of the gene or mRNA can be used in an antisense approach to inhibit translation of endogenous mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon.

Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but can be used in accordance with the invention. Whether designed to hybridize to the 5', 3', or coding region of an mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in

length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides, or at least 50 nucleotides.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (as described, e.g., in Letsinger et al., *Proc. Natl. Acad. Sci. USA* 86:6553, 1989; Lemaitre et al., *Proc. Natl. Acad. Sci. USA* 84:648, 1987; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, for example, PCT Publication No. WO 89/10134), or hybridization-triggered cleavage agents (see, for example, Krol et al., *BioTechniques* 6:958, 1988), or intercalating agents (see, for example, Zon, *Pharm. Res.* 5:539, 1988). To this end, the oligonucleotide can be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethyl-aminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil,

beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-theouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 2-(3-amino-3-N-2-carboxypropl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal, or an analog of any of these backbones.

In yet another embodiment, the antisense oligonucleotide is an α-anomeric oligonucleotide. An α-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gautier et al., *Nucl. Acids. Res.* 15:6625, 1987). The oligonucleotide is a 2'-0-methylribonucleotide (Inoue et al., *Nucl. Acids Res.* 15:6131, 1987), or a chimeric RNA-DNA analog (Inoue et al., *FEBS Lett.* 215:327, 1987).

Antisense oligonucleotides of the invention can be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (*Nucl. Acids Res.* 16:3209, 1988), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., *Proc. Natl. Acad. Sci. USA* 85:7448, 1988).

The antisense molecules should be delivered to cells that express nucleic acids or polypeptides of the invention *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular concentrations of the antisense molecule sufficient to suppress translation of endogenous mRNAs. Therefore, a preferred approach uses a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a

strong *pol* III or *pol* II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous transcripts of nucleic acids of the invention and thereby prevent translation of the endogenous mRNA. For example, a vector can be introduced *in vivo* such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA.

Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to: the SV40 early promoter region (Bernoist et al., *Nature* 290:304, 1981); the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell* 22:787-797, 1988); the herpes thymidine kinase promoter (Wagner et al., *Proc. Natl. Acad. Sci. USA* 78:1441, 1981); or the regulatory sequences of the metallothionein gene (Brinster et al., *Nature* 296:39, 1988).

Ribozymes

Ribozyme molecules designed to catalytically cleave mRNA transcripts of nucleic acids of the invention can be used to prevent translation and expression of mRNA of the invention. (see, e.g., PCT Publication WO 90/11364; Saraver et al., *Science* 247:1222, 1990). While various ribozymes that cleave mRNA at site-specific recognition sequences can be used to destroy mRNAs of the invention, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art (Haseloff et al., *Nature* 334:585, 1988). There are numerous examples of potential hammerhead ribozyme cleavage sites within the nucleotide sequence of human cDNA of the invention. Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the mRNA, i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes"), such as the one that occurs naturally in *Tetrahymena thermophila* (known as the IVS or L-19 IVS RNA), and which has been extensively described by Cech and his collaborators (Zaug et al., *Science* 224:574, 1984; Zaug et al., *Science*, 231:470, 1986; Zug et al., *Nature* 324:429, 1986; PCT Application No. WO 88/04300; and Been et al., *Cell* 47:207, 1986). The Cech-type ribozymes have an eight base-pair sequence that hybridizes to a target RNA sequence, whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes that target eight base-pair active site sequences present in nucleic acids of the invention.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.), and should be delivered to cells which express nucleic acids or polypeptides of the invention *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive *pol* III or *pol* II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Other Methods for Modulating IFT 20, 27, 46, 52, 57, 72, 88, 122, and 139, and Che-2 Expression

Endogenous expression of a gene of the invention can also be modulated by inactivating the endogenous gene or its promoter using targeted homologous recombination (see, e.g., U.S. Patent No. 5,464,764). For example, a mutant, non-functional gene of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous gene of the invention (either the coding regions or regulatory regions of the gene of the invention) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the endogenous gene of the invention *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the gene of the invention. Such approaches are particularly suited for use in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive gene of the invention. However, this approach can be adapted for use in humans, provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Alternatively, endogenous expression of a gene of the invention can be modulated by targeting deoxyribonucleotide sequences complementary to the regulatory region of the gene of the invention (*i.e.*, the promoter and/or enhancers of a gene of the invention) to form triple helical structures that prevent transcription of the gene of the invention in target cells in the body (Helene, *Anticancer Drug Res.* 6:569, 1981; Helene et al., *Ann. N.Y. Acad. Sci.* 660:27, 1992; and Maher, *Bioassays* 14:807, 1992).

The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid of the invention. Such methods comprise formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid of the invention. Such compositions can further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent that modulates expression or activity of a polypeptide or nucleic acid of the invention and one or more additional active compounds.

The agent that modulates expression or activity can, for example, be a small molecule. For example, such small molecules include peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

It is understood that appropriate doses of small molecule agents and protein or polypeptide agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of these agents will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the agent to have upon the nucleic acid or polypeptide of the invention. Examples of doses of a small molecule include milligram or microgram amounts per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100

micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). Examples of doses of a protein or polypeptide include gram, milligram or microgram amounts per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 5 grams per kilogram, about 100 micrograms per kilogram to about 500 milligrams per kilogram, or about 1 milligram per kilogram to about 50 milligrams per kilogram). For antibodies, examples of dosages are from about 0.1 milligram per kilogram to 100 milligrams per kilogram of body weight (generally 10 milligrams per kilogram to 20 milligrams per kilogram). If the antibody is to act in the brain, a dosage of 50 milligrams per kilogram to 100 milligrams per kilogram is usually appropriate. It is furthermore understood that appropriate doses of one of these agents depend upon the potency of the agent with respect to the expression or activity to be modulated. Such appropriate doses can be determined using the assays described herein. When one or more of these agents is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher can, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific agent employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

As an alternative to making determinations based on the absolute expression level of selected genes, determinations may be based on the normalized expression levels of these genes. Expression levels are normalized by correcting the absolute expression level of a gene encoding a polypeptide of the invention by comparing its expression to the expression of a different gene, e.g., a housekeeping gene that is constitutively expressed. Suitable genes for normalization include housekeeping genes such as the actin gene. This normalization allows the comparison of the expression level in one sample (e.g., a patient sample), to another sample, or between samples from different sources.

Alternatively, the expression level can be provided as a relative expression level. To determine a relative expression level of a gene, the level of expression of the gene is determined for 10 or more samples of different endothelial (e.g. intestinal endothelium, airway endothelium, or other mucosal epithelium) cell isolates, preferably 50 or more samples, prior to the determination of the expression level for the sample in question. The mean expression level of each of the genes

assayed in the larger number of samples is determined and this is used as a baseline expression level for the gene(s) in question. The expression level of the gene determined for the test sample (absolute level of expression) is then divided by the mean expression value obtained for that gene. This provides a relative expression level and aids in identifying extreme cases of disorders associated with aberrant expression of a gene encoding a polypeptide of the invention protein or with aberrant expression of a ligand thereof.

Preferably, the samples used in the baseline determination will be from either or both of cells which aberrantly express a gene encoding a polypeptide of the invention or a ligand thereof (i.e. 'diseased cells') and cells which express a gene encoding a polypeptide of the invention at a normal levelor a ligand thereof (i.e. 'normal' cells). The choice of the cell source is dependent on the use of the relative expression level. Using expression found in normal tissues as a mean expression score aids in validating whether aberrance in expression of a gene encoding a polypeptide of the invention occurs specifically in diseased cells. Such a use is particularly important in identifying whether a gene encoding a polypeptide of the invention can serve as a target gene. In addition, as more data is accumulated, the mean expression value can be revised, providing improved relative expression values based on accumulated data. Expression data from endothelial cells (e.g. mucosal endothelial cells) provides a means for grading the severity of the disorder.

Detecting Proteins Associated with IFT 20, 27, 46, 52, 57, 72, 88, 122, or 139, or Che-2

The invention also features polypeptides that interact with (e.g., bind directly or indirectly) IFT 20, 27, 46, 52, 57, 72, 88, 122, or 139, or Che-2. Any method suitable for detecting protein-protein interactions may be employed for identifying transmembrane proteins, intracellular, or extracellular proteins that interact with polypeptides of the invention. Among the traditional methods which may be employed are co-immunoprecipitation, cross-linking and co-purification through gradients or chromatographic columns of cell lysates or proteins obtained from cell lysates and the use of polypeptides of the invention to identify proteins in the lysate that interact with polypeptides of the invention. For these assays, the polypeptide of the invention can be full length polypeptide of the invention, a soluble extracellular domain of a polypeptide of the invention, or some other suitable polypeptide of the invention. Once isolated, such an interacting protein can be identified and cloned and then used, in conjunction with standard techniques, to identify proteins with which it interacts. For example, at least a portion of the amino acid sequence of a protein

which interacts with the polypeptide of the invention can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique. The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding the interacting protein. Screening may be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (Ausubel, *supra;* and "PCR Protocols: A Guide to Methods and Applications," Innis et al., eds. Academic Press, Inc., NY, 1990).

Additionally, methods may be employed which result directly in the identification of genes which encode proteins which interact with polypeptides of the invention. These methods include, for example, screening expression libraries, in a manner similar to the well known technique of antibody probing of λgt11 libraries, using labeled polypeptide of the invention or a fusion protein of the invention, e.g., a polypeptide of the invention or domain thereof fused to a marker such as an enzyme, fluorescent dye, a luminescent protein, or to an IgFc domain.

There are also methods capable of detecting protein interaction. A method that detects protein interactions *in vivo* is the two-hybrid system (Chien et al., *Proc. Natl. Acad. Sci. USA*, 88:9578, 1991). A kit for practicing this method is available from Clontech (Palo Alto, CA).

III. Uses

A. Human and Animal Therapeutic Uses

Inhibition of IFT Particle Protein Function (Spermiogenesis and Contraception)

The sperm flagellum consists of a 9 + 2 axoneme which generates sperm movement. In addition, the flagellum of the mammalian sperm contains accessory structures, viz. the fibrous sheath and the outer dense fibers, which are believed to stiffen the flagellum and enable it to function in the viscous environment of the female reproductive tract (Witman, Introduction to cilia and flagella. *In*: *Ciliary and Flagellar Membranes* (ed. R. A. Bloodgood). Plenum Press, NY. pp. 1-30, 1990). During spermiogenesis in the mammalian testis, a 9 + 2 axoneme is formed within a flagellar membrane; as in other flagella, proper assembly of this axoneme must require IFT to move the axonemal precursors to the tip of the flagellum, where they are assembled. After assembly of the axoneme, the accessory structures are assembled *beginning at the distal tip of the flagellum and progressing toward its base* (Oko and Clermont, Mammalian spermatozoa: structure and assembly of the tail. *In*: *Controls of Sperm Motility: Biological and*

Clinical Aspects. (ed. C. Gagnon), CRC Press, Boca Raton, Florida, pp. 3-27, 1990). Concomitant with this assembly, accessory structure proteins, previously present in the cytoplasm of the spermiogenic cells, are mobilized from the cytoplasm to the flagellum. This transport of outer dense fiber and fibrous sheath precursors from the cytoplasm to the tip of the forming flagellum also must be dependent on IFT. Indeed, in the Tg737 insertional mutant mouse, which has a defect in IFT, the sperm flagellum is not formed, and accessory structure precursors remain in the cell body (J. San Agustin, G. Pazour and G. Witman, unpublished results). Therefore, in humans, a defect in IFT would be expected to result in azoospermia and oligozoospermia. Azoospermia and oligospermia are among the most common reasons for infertility in men. It is likely that a genetic defect in an IFT protein is responsible for at least some cases of azoospermia and oligozoospermia.

Because IFT is essential for sperm flagella formation, a drug that inhibits IFT would block spermiogenesis and prevent fertility. Ideally, this drug would 1) target an IFT protein or isoform expressed specifically in the testis, and 2) be able to cross the blood/testis barrier. Such a drug, which can be administered to men orally or by any other preferred route, would be specific for IFT in the testis and can serve as a male contraceptive.

Human and animal parasites

Many human and animal eukaryotic parasites have cilia or flagella that are critical for their ability to infect and survive in their hosts. Therefore, these parasites are likely to be susceptible to anti-IFT drugs that would block formation and function of their cilia or flagella. Ideally, anti-IFT drugs would be tailored to bind specifically to the parasite's IFT proteins in order to avoid undesirable effects on the host's cilia and flagella. However, because short-term drug-induced inhibition of IFT in humans and animals is unlikely to cause long-lasting side effects, the host most likely can tolerate treatment with non-host specific anti-IFT drugs for the short time periods necessary to eliminate a parasite. Some examples of ciliated or flagellated parasites that would be amenable to treatment by anti-IFT drugs are given below.

Parasitic protozoa

Parasitic protozoa cause numerous human and animal diseases, including malaria, African sleeping sickness, trypanosomosis, leishmanioses, trichomonosis, and giardiosis. All of

these parasites have stages of their life cycles that are dependent on cilia or flagella, and in many cases cilia or flagella are present during those life-cycle stages that occur in the human or animal host, making the parasites vulnerable to drugs which would block IFT and hence inhibit cilia or flagella assembly or function.

One example is *Giardia sp.*, an intestinal parasite which causes debilitating diarrhea and other symptoms in humans. *Giardia* live primarily in the upper portion of the lumen of the small intestine, where they attach to the enterocytes of the intestinal wall (Marquardt et al., *Parasitology and Vector Biology*, 2nd ed., Academic Press, San Diego, 2000). Each *Giardia* cell has 8 motile flagella, which presumably are used to stay in the upper portion of the small intestine and to reach the site of attachment. An anti-IFT drug, taken orally, would inhibit assembly of flagella in newly divided *Giardia* and cause disassembly of previously formed flagella in non-dividing *Giardia*. As a result, the *Giardia* would not be able to move up the lumen of the small intestine, or reach the wall of the small intestine to attach to its surface. In the absence of motility and anchorage, the *Giardia* would be passed out of the intestine and the infection would be eliminated.

An anti-IFT drug also would be effective against trypanosomes, which are responsible for diseases such as African sleeping sickness and Chagas' disease in humans, and nagana in cattle. Trypanosomes circulate in the blood where they reproduce by asexual division. These parasites are characterized by the presence of a single motile flagellum that arises from a flagellar pocket and is enclosed by a sheath called the undulating membrane. An anti-IFT drug would block assembly of the flagellum and the flagellar sheath and affect the trypanosome's life cycle and host-parasite interactions in at least three ways. First, by blocking flagellar assembly, it would affect those normal life processes that are dependent upon flagellar motility. Second, the ability of the trypanosome to evade the human or animal host's defense mechanisms in the bloodstream is dependent on its production of a protective glycoprotein coat that covers both the cell body and the flagellar sheath (Marquandt et al., 1999). Movement of these proteins onto the flagellar sheath is dependent upon IFT (Rosenbaum and Witman, 2001; Bloodgood, 2000). In the absence of IFT, the protective coating would not be present on the flagellar sheath and the parasite would be susceptible to attack by the host immune system. Third, the flagellum is essential for the trypanosome's attachment to, and infection of, the insect vector (e.g., tsetse fly, kissing bugs) which take up the parasite from the human or animal bloodstream and then spread

the disease to other humans or animals (Vickerman and Tetley, Flagellar surfaces of parasitic protozoa and their role in attachment, *In: Ciliary and Flagellar Membranes* (ed. R.A. Bloodgood), Plenum Press, N.Y., 1989; Marquardt et al., *Parasitology and Vector Biology*, 2nd ed., Academic Press, San Diego, 2000). By eliminating the trypanosome's flagellum prior to uptake of the parasite by the insect vector, the trypanosome's life cycle would be interrupted and transmission of the parasite to new hosts would be prevented. Because trypanosomes live in the bloodstream, they would be very susceptible to anti-IFT drugs administered by intravenous injection.

Trichomonads are flagellated parasitic protozoans that likewise can be treated with an anti-IFT drug. *Tritrichomonas foetus* causes trichomonad abortion in cattle and other bovines. In the United States alone, it is estimated that there is an annual loss of $650 million to the cattle industry from trichomonosis (Marquardt et al., *Parasitology and Vector Biology*, 2nd ed., Academic Press, San Diego, 2000). *T. foetus* infects the reproductive tracts of both cows and bulls. It reproduces by asexual division, and is spread by sexual intercourse. Upon introduction into the reproductive tract of a cow during sexual intercourse, the trichomonads reproduce in the vagina. If the animal becomes pregnant, the organisms may invade the uterus and infect the developing fetus, causing abortion (Marquardt et al., *Parasitology and Vector Biology*, 2nd ed., Academic Press, San Diego, 2000). Trichomonads have three or more motile anterior flagella, and a motile recurrent flagellum usually attached to the body by an undulating membrane. These flagella presumably are essential for movement of the parasite up the reproductive tract. Treatment of infected cows with an anti-IFT drug delivered orally or by intravagina suppository would prevent assembly of the flagella of newly divided cells, and result in loss of flagella of non-dividing cells. In the absence of the flagellum, the trichomonads would be immotile and would be unable to move up the female reproductive tract and cause abortion. Moreover, rendering the parasites immotile and paralyzed would decrease their ability to withstand the host's own immune defenses, so that the infection can be completely eliminated. Similarly, treatment of bulls with anti-IFT drugs administered orally would render the trichomonads immotile, so they can be eliminated by the bull's own immune system.

Trichomonas vaginalis causes trichomonas vaginitis in humans and can be similarly treated to eliminate infection.

Parasitic Platyhelminthes

Many human and animal parasites are members of the phylum Platyhelminthes. Parasitic Platyhelminthes include liver flukes, intestinal flukes, lung flukes and other trematodes, and cestodes or tapeworms. In humans, these parasites cause severe illnesses such as schistosomosis, dicrocoeliosis, clonorchiasis, opisthorchosis, echinostomatiasis, heterophyidosis, swimmer's itch, taeniosis, and cysticercosis. Infection of livestock by these parasites results in huge economic losses in cattle- and sheep-raising areas. All Platyhelminthes have an excretory system based on the flame cell or protonephridium, in which currents are created by a tuft of vigorously beating cilia (Bogitsh and Cheng, *Human Parisitology*, Academic Press, San Diego, 1979; Marquardt et al., *Parasitology and Vector Biology*, 2nd ed., Academic Press, San Diego, 2000). The flame cell cilia may have additional sensory or osmoregulatory functions. In addition, some parasitic Platyhelminthes, such as flatworms, have external cilia that are believed to have a sensory role (Marquardt et al., *Parasitology and Vector Biology*, 2nd ed., Academic Press, San Diego, 2000). Because all cilia and flagella appear to be dependent upon IFT for their assembly and maintenance, one would expect that a drug that inhibited IFT would prevent assembly of the cilia, or result in disassembly of previously formed cilia, in these parasites, causing malfunction of the parasite's osmoregulatory and/or nervous systems. Thus, treatment of infected humans or animals with an anti-IFT drug would control or cure infections by these Platyhelminthes. The anti-IFT drug would be administered orally or intravenously, depending on the site of infection.

Parasitic nematodes

Over half of the world's population is estimated to be infected by at least one or more species of parasitic nematodes. Animal parasitic nematodes are also widespread and exhibit a wide diversity. Nematodes sense their environment with a set of sensory neurons in which the dendrites are ciliated. These sensory cilia extend outside of the tough nematode cuticle and are exposed to the outer environment where they are responsible for detecting various chemical signals as well as monitor osmotic conditions. Nematodes, for example, will move away from high salt conditions so as to protect themselves from osmotic stress. In nematode mutants where the sensory cilia are structurally defective, the organism is unable to identify dangerous levels of salt; these mutants are also defective in chemosensation (Perkins et al., *Dev. Biol.*, 117:456-487, 1986). Since IFT is required for the formation and maintenance of these sensory cilia, blocking

IFT with various anti-IFT agents should result in nematodes that have lost their ability to sense their environment. This will adversely affect the organism's ability to determine where it is going and in finding sexual partners. As described in detail under Combating phytopathogenic nematodes, nematode-specific RNA interference should also work in combating human and animal nematode parasites.

Ascariasis, caused by *Ascaris lumbricoides* is globally distributed with more than 1.4 billion persons infected throughout the world (Khuroo, *Gastroenterol Clin North Am*, 25:553-577 1996). *Ascaris* infection begins with the ingestion of infective eggs. Once in the small intestine, the eggs hatch and the resulting juveniles pass through the intestinal wall and enter the bloodstream and eventually end up in the lungs. After entering the airspace, *Ascaris* move up into the pharynx where they are swallowed and reappear in the small intestine. The adult female can lay as many as 200,000 eggs a day in the small intestine. Potentially fatal pathologies includes ascariasis pneumonia due to pulmonary hemorrhaging and inflammation and physical blockage of the gastrointestinal tract due to a large mass of the *Ascaris* nematodes. *Ascaris suum* is a common pig parasite that also infects humans but the juvenile gets no farther than the lungs; pathology includes ascariasis pneumonia. In the intestine, *Ascaris spp* can be susceptible to orally administered anti-IFT agents. Anti-IFT agents can also be administered by intra-venous injection to treat juvenile nematodes in the bloodstream.

The nematodes known as hookworms infect over a billion people. Most of these infections are caused by either *Ancylostoma duodenale* or *Necator americanus*. The life cycle of hookworms is similar to that of *Ascaris* and suggests that anti-IFT treatments that are successful in treating ascariasis should also be successful in treating hookworm infections.

The nematodes known as whipworms are also estimated to infect over a billion people. These infections, caused by *Trichuris trichiura,* are typically confined to the large intestine. Anti-IFT agents can be administered either orally or by suppository.

Animals are infected at a high rate by a large and diverse group of nematodes. In all cases, however, the nematodes may be susceptible to anti-IFT treatments described above. These include nematode-specific RNA interference directed against expression of IFT subunits. The appropriate IFT-derived double stranded RNA can be administered directly to the infected animals or used as a preventive treatment prior to ingestion of the infective eggs. Alternatively, transgenic animal strains can be developed that contain stable transformed DNA that encodes

double stranded RNA derived from nematode-specific sequences that encode IFT proteins. In this way, entire populations of livestock can be produced that will no longer support infection by one or more species of parasitic nematode. Two obvious benefits of this approach are (1) that animals will not have to be periodically treated to reduce or remove infection and (2) it may be possible to eradicate specific infectious species by making all hosts resistant.

Restoration of IFT Particle Protein Function

Gene Therapy

The gene constructs of the invention can also be used as a part of a gene therapy protocol to deliver nucleic acids encoding either an agonistic or antagonistic form of a polypeptide of the invention. The invention features expression vectors for *in vivo* transfection and expression of a polypeptide of the invention in particular cell types so as to reconstitute the function of, or alternatively, antagonize the function of a polypeptide of the invention in a cell in which that polypeptide is misexpressed. Expression constructs of polypeptides of the invention, may be administered in any biologically effective carrier, e.g. any formulation or composition capable of effectively delivering a gene of the invention to cells *in vivo*. Approaches include insertion of the subject gene in viral vectors including recombinant retroviruses, adenovirus, adeno-associated virus, and herpes simplex virus-1, or recombinant bacterial or eukaryotic plasmids. Viral vectors transfect cells directly; plasmid DNA can be delivered with the help of, for example, cationic liposomes (lipofectin) or derivatized (e.g. antibody conjugated), polylysine conjugates, gramacidin S, artificial viral envelopes or other such intracellular carriers, as well as direct injection of the gene construct or $CaPO_4$ precipitation carried out *in vivo*.

A preferred approach for *in vivo* introduction of nucleic acid into a cell is by use of a viral vector containing nucleic acid, e.g. a cDNA, encoding a polypeptide of the invention. Infection of cells with a viral vector has the advantage that a large proportion of the targeted cells can receive the nucleic acid. Additionally, molecules encoded within the viral vector, e.g., by a cDNA contained in the viral vector, are expressed efficiently in cells which have taken up viral vector nucleic acid.

Retrovirus vectors and adeno-associated virus vectors can be used as a recombinant gene delivery system for the transfer of exogenous genes *in vivo*, particularly into humans. These vectors provide efficient delivery of genes into cells, and the transferred nucleic acids are stably

integrated into the chromosomal DNA of the host. The development of specialized cell lines (termed "packaging cells") which produce only replication-defective retroviruses has increased the utility of retroviruses for gene therapy, and defective retroviruses are characterized for use in gene transfer for gene therapy purposes (for a review see Miller, A.D. (1990) *Blood* 76:271). A replication defective retrovirus can be packaged into virions which can be used to infect a target cell through the use of a helper virus by standard techniques. Protocols for producing recombinant retroviruses and for infecting cells *in vitro* or *in vivo* with such viruses can be found in Current Protocols in Molecular Biology, Ausubel, F.M. et al. (eds.) Greene Publishing Associates, (1989), Sections 9.10-9.14 and other standard laboratory manuals. Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are known to those skilled in the art. Examples of suitable packaging virus lines for preparing both ecotropic and amphotropic retroviral systems include yCrip, yCre, y2 and yAm. Retroviruses have been used to introduce a variety of genes into many different cell types, including epithelial cells, *in vitro* and/or *in vivo* (see for example Eglitis, et al. (1985) *Science* 230:1395-1398; Danos and Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:6460-6464; Wilson et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:3014-3018; Armentano et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6141-6145; Huber et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8039-8043; Ferry et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8377-8381; Chowdhury et al. (1991) *Science* 254:1802-1805; van Beusechem et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:7640-7644; Kay et al. (1992) *Human Gene Therapy* 3:641-647; Dai et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:10892-10895; Hwu et al. (1993) *J. Immunol.* 150:4104-4115; U.S. Patent No. 4,868,116; U.S. Patent No. 4,980,286; PCT Application WO 89/07136; PCT Application WO 89/02468; PCT Application WO 89/05345; and PCT Application WO 92/07573).

Another viral gene delivery system useful in the present invention utilizes adenovirus-derived vectors. The genome of an adenovirus can be manipulated such that it encodes and expresses a gene product of interest but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. See, for example, Berkner et al. (1988) *BioTechniques* 6:616; Rosenfeld et al. (1991) *Science* 252:431-434; and Rosenfeld et al. (1992) *Cell* 68:143-155. Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other strains of adenovirus (e.g., Ad2, Ad3, Ad7 etc.) are known to those skilled in the art. Recombinant adenoviruses can be advantageous in certain circumstances in that they are not capable of

infecting nondividing cells and can be used to infect a wide variety of cell types, including epithelial cells (Rosenfeld et al. (1992) cited *supra*). Furthermore, the virus particle is relatively stable and amenable to purification and concentration, and as above, can be modified so as to affect the spectrum of infectivity. Additionally, introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal, thereby avoiding potential problems that can occur as a result of insertional mutagenesis *in situ* where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA). Moreover, the carrying capacity of the adenoviral genome for foreign DNA is large (up to 8 kilobases) relative to other gene delivery vectors (Berkner et al. cited *supra*; Haj-Ahmand and Graham (1986) *J. Virol.* 57:267).

Yet another viral vector system useful for delivery of the subject gene is the adeno-associated virus (AAV). Adeno-associated virus is a naturally occurring defective virus that requires another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient replication and a productive life cycle. (For a review see Muzyczka et al. (1992) *Curr. Topics in Micro. and Immunol.*158:97-129). It is also one of the few viruses that may integrate its DNA into non-dividing cells, and exhibits a high frequency of stable integration (see for example Flotte et al. (1992) *Am. J. Respir. Cell. Mol. Biol.* 7:349-356; Samulski et al. (1989) *J. Virol.* 63:3822-3828; and McLaughlin et al. (1989) *J. Virol.* 62:1963-1973). Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate. Space for exogenous DNA is limited to about 4.5 kb. An AAV vector such as that described in Tratschin et al. (1985) *Mol. Cell. Biol.* 5:3251-3260 can be used to introduce DNA into cells. A variety of nucleic acids have been introduced into different cell types using AAV vectors (see for example Hermonat et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6466-6470; Tratschin et al. (1985) *Mol. Cell. Biol.* 4:2072-2081; Wondisford et al. (1988) *Mol. Endocrinol.* 2:32-39; Tratschin et al. (1984) *J. Virol.* 51:611-619; and Flotte et al. (1993) *J. Biol. Chem.* 268:3781-3790).

In addition to viral transfer methods, such as those illustrated above, non-viral methods can also be employed to cause expression of a polypeptide of the invention in the tissue of an animal. Most nonviral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral gene delivery systems of the present invention rely on endocytic pathways for the uptake of the subject gene of the invention by the targeted cell. Exemplary gene delivery

systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes. Other embodiments include plasmid injection systems such as are described in Meuli et al. (2001) *J Invest Dermatol.* 116(1):131-135; Cohen et al. (2000) *Gene Ther* 7(22):1896-905; or Tam et al. (2000) *Gene Ther* 7(21):1867-74.

In a representative embodiment, a gene encoding a polypeptide of the invention can be entrapped in liposomes bearing positive charges on their surface (e.g., lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno et al. (1992) *No Shinkei Geka* 20:547-551; PCT publication WO91/06309; Japanese patent application 1047381; and European patent publication EP-A-43075).

In clinical settings, the gene delivery systems for the therapeutic gene of the invention can be introduced into a patient by any of a number of methods, each of which is familiar in the art. For instance, a pharmaceutical preparation of the gene delivery system can be introduced systemically, e.g. by intravenous injection, and specific transduction of the protein in the target cells occurs predominantly from specificity of transfection provided by the gene delivery vehicle, cell-type or tissue-type expression due to the transcriptional regulatory sequences controlling expression of the receptor gene, or a combination thereof. In other embodiments, initial delivery of the recombinant gene is more limited with introduction into the animal being quite localized. For example, the gene delivery vehicle can be introduced by catheter (see U.S. Patent 5,328,470) or by stereotactic injection (e.g. Chen et al. (1994) *PNAS* 91: 3054-3057).

The pharmaceutical preparation of the gene therapy construct can consist essentially of the gene delivery system in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery system can be produced in tact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can comprise one or more cells which produce the gene delivery system.

Compounds that affect the activity of IFT

Drugs that stop IFT can have important anti-fertility, anti-parasitic, and anti-pesticide activities while drugs that enhance IFT can be used to improve ciliary function in patients with diseases caused by reduced IFT. *In vitro* binding assays can be used to identify compounds that bind with high affinity to IFT particle proteins. The identified compounds can then be used in *in*

vivo assays to determine if they affect ciliary assembly. Compounds may include small organic or inorganic molecules, peptides, peptidomimetics, nucleic acids, or carbohydrates.

Genes encoding IFT particle proteins from humans or other species can be expressed in bacterial, mammalian, insect, or other cells and purified. These proteins can then be used in high-throughput screens of chemical compounds to identify those that bind, e.g., with high affinity. The identified compounds can then be screened to determine if they affect ciliary assembly.

The effect of the drugs on ciliary assembly can be determined by adding them to cultures of ciliated organisms and observing how this affects the cilia. For example, the effect on trypanosome ciliary assembly can be determined by adding the compounds to axenic cultures and observing the cells by light microscopy to see if ciliary assembly is affected (Tyler and Engman, *Cell Motil Cytoskeleton*, 46(4):269-78, 2000). The effect on nemotode sensory cilia can be determined by assaying the ability of treated worms to take up membrane-soluble fluorescent dyes. Nematodes that lack cilia do not take up the dyes (Starich et al, *Genetics*, 139:171-188, 1995). The effect on mammalian primary cilia assembly can be determined by adding the compounds to cultured mammalian cells and assaying ciliary assembly by immunofluorescence microscopy (Wheatley et al., *Cell Biol. Int.*, 20:73-81, 1996).

Compounds that bind to IFT particle proteins can also be screened to determine if they enhance IFT. This can be done using cultured cells derived from the Tg737 mutant mouse (Taulman et al., *Mol Biol Cell*, 12(3):589-99, 2001). This mouse has a reduced amount of the Tg737/IFT88 protein (Taulman et al., *Mol Biol Cell*,12(3):589-99, 2001; Pazour unpublished observation) and shows defects in assembly of kidney primary cilia (Pazour et al., *Mol. Biol. Cell*, 151:709-718, 2000). Presumably the cells derived from this mouse will also show defects in ciliary assembly and can be used to screen for compounds that enhance IFT to compensate for the reduced amount of protein.

Assays for Compounds Capable of Restoring or Inhibiting IFT Function

Even after a particular intracellular target is selected, the means by which new IFT protein function restoration agents are identified pose certain challenges. Despite the increased use of rational drug design, a preferred method continues to be the mass screening of compound "libraries" for active agents by exposing cultures of cells with cilia or flagella to the test

compounds and assaying for inhibition or restoration of normal IFT protein activity. In testing thousands or tens of thousands of compounds, however, a correspondingly large number of cell cultures of interest must be grown over time periods which are relatively long. Moreover, a compound that is found to inhibit or restore normal IFT protein function in culture may be acting not on the desired target but on a different, less unique component of the IFT system, with the result that the compound may act against host cells as well and thereby produce unacceptable side effects. Consequently, there is a need for an assay or screening method that more specifically identifies those agents that are active against a certain intracellular target. Additionally, there is a need for assay methods having greater throughput, that is, assay methods that reduce the time and materials needed to test each compound of interest.

The invention provides methods for identifying agents capable of modululating, e.g., restoring or inhibiting, IFT function. In some methods, screening for potential or candidate IFT function restorative or inhibitory agents is accomplished by identifying those compounds (e.g., small organic molecules) that interact with, e.g., bind directly or indirectly, to an IFT polypeptide and/or inhibit the activity of a IFT polypeptide or the expression of an IFT gene. For example, screens can be performed that identify those compounds that inhibit an IFT activity described herein.

In various suitable methods, screening for anti-IFT agents can be accomplished by (i) identifying those compounds that bind to IFT (and are thus candidate anti-IFT compounds) and (ii) further testing such candidate compounds for their ability to inhibit intraflagellar transport in vitro or in vivo, in which case they are anti-IFT agents.

Specific binding of a test compound to a polypeptide can be detected, for example, in vitro by reversibly or irreversibly immobilizing the test compound(s) on a substrate, e.g., the surface of a well of a 96-well polystyrene microtitre plate. Methods for immobilizing polypeptides and other small molecules are well known in the art. For example, the microtitre plates can be coated with a IFT polypeptide by adding the polypeptide in a solution (typically, at a concentration of 0.05 to 1 mg/ml in a volume of 1-100 ml) to each well, and incubating the plates at room temperature to 37°C for 0.1 to 36 hours. Polypeptides that are not bound to the plate can be removed by shaking the excess solution from the plate, and then washing the plate (once or repeatedly) with water or a buffer. Typically, the polypeptide is in water or a buffer. The plate is then washed with a buffer that lacks the bound polypeptide. To block the free

protein-binding sites on the plates, the plates are blocked with a protein that is unrelated to the bound polypeptide. For example, 300 ml of bovine serum albumin (BSA) at a concentration of 2 mg/ml in Tris-HCl is suitable. Suitable substrates include those substrates that contain a defined cross-linking chemistry (e.g., plastic substrates, such as polystyrene, styrene, or polypropylene substrates from Corning Costar Corp., Cambridge, MA, for example). If desired, a beaded particle, e.g., beaded agarose or beaded Sepharose, can be used as the substrate. The IFT is then added to the coated plate and allowed to bind to the test compound (e.g., at 37°C for 0.5-12 hours). The plate then is rinsed as described above.

Binding of the test compound to the IFT can be detected by any of a variety of known methods. For example, an antibody that specifically binds to a IFT polypeptide can be used in an immunoassay. If desired, the antibody can be labeled (e.g., fluorescently or with a radioisotope) and detected directly (see, e.g., West and McMahon, J. Cell Biol. 74:264, 1977). Alternatively, a second antibody can be used for detection (e.g., a labeled antibody that binds to the Fc portion of an anti-YphC antibody). In an alternative detection method, the IFT polypeptide is labeled, and the label is detected (e.g., by labeling a IFT polypeptide with a radioisotope, fluorophore, chromophore, or the like). In still another method, the IFT polypeptide is produced as a fusion protein with a protein that can be detected optically, e.g., using green fluorescent protein (which can be detected under UV light). In an alternative method, the polypeptide can be produced as a fusion protein with an enzyme having a detectable enzymatic activity, such as horseradish peroxidase, alkaline phosphatase, J-galactosidase, or glucose oxidase. Genes encoding all of these enzymes have been cloned and are available for use by those of skill in the art. If desired, the fusion protein can include an antigen, and such an antigen can be detected and measured with a polyclonal or monoclonal antibody using conventional methods. Suitable antigens include enzymes (e.g., horseradish peroxidase, alkaline phosphatase, and J-galactosidase) and non-enzymatic polypeptides (e.g., serum proteins, such as BSA and globulins, and milk proteins, such as caseins).

In various in vivo methods for identifying polypeptides that bind to IFT, the conventional two-hybrid assays of protein/protein interactions can be used (see e.g., Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578, 1991; Fields et al., U.S. Pat. No. 5,283,173; Fields and Song, Nature, 340:245, 1989; Le Douarin et al., Nucleic Acids Research, 23:876, 1995; Vidal et al., Proc. Natl. Acad. Sci. USA, 93:10315-10320, 1996; and White, Proc. Natl. Acad. Sci. USA, 93:10001-

10003, 1996). Generally, the two-hybrid methods involve in vivo reconstitution of two separable domains of a transcription factor. One fusion protein contains the IFT polypeptide fused to either a transactivator domain or DNA binding domain of a transcription factor (e.g., of Gal4). The other fusion protein contains a test polypeptide fused to either the DNA binding domain or a transactivator domain of a transcription factor. Once brought together in a single cell (e.g., a yeast cell or mammalian cell), one of the fusion proteins contains the transactivator domain and the other fusion protein contains the DNA binding domain. Therefore, binding of the IFT polypeptide to the test polypeptide (i.e., candidate anti-IFT agent) reconstitutes the transcription factor. Reconstitution of the transcription factor can be detected by detecting expression of a gene (i.e., a reporter gene) that is operably linked to a DNA sequence that is bound by the DNA binding domain of the transcription factor. Kits for practicing various two-hybrid methods are commercially available (e.g., from Clontech; Palo Alto, CA).

The methods described above can be used for high throughput screening of numerous test compounds to identify candidate compounds that modulate the function or activity of IFT particle proteins. Having identified a test compound as a candidate compound, the candidate compound can be further tested for inhibition of intraflagellar transport in vitro or in vivo (e.g., using a cell, animal, e.g., rodent, model) if desired.

In vitro, further testing can be accomplished by means known to those in the art such as an enzyme inhibition assay or a whole-cell growth inhibition assay. For example, an agar dilution assay identifies a substance that inhibits cell growth. Microtiter plates are prepared with serial dilutions of the test compound, adding to the preparation a given amount of growth substrate, and providing a preparation of cells. Inhibition of cell growth is determined, for example, by observing changes in optical densities of the cell cultures.

Inhibition of cell growth is demonstrated, for example, by comparing (in the presence and absence of a test compound) the rate of growth or the absolute growth of cells. Inhibition includes a reduction in the rate of growth or absolute growth by at least 20%. Particularly potent test compounds can further reduce the growth rate (e.g., by at least 25%, 30%, 40%, 50%, 75%, 80%, or 90%).

Animal (e.g., rodent such as murine) models of intraflagellar transport are known to those of skill in the art, and such animal model systems are acceptable for screening potential compounds capable of restoring intraflagellar transport function as an indication of their

therapeutic efficacy in human patients. In a typical in vivo assay, an animal is treated with a compound that is capable of restoring IFT function, and conventional methods and criteria are used to diagnose the mammal as having impaired intraflagellar transport. The candidate IFT restorative agent then is administered to the mammal at a dosage of 1-100 mg/kg of body weight, and the mammal is monitored for signs of IFT function restoration. Of course, the results obtained in the presence of the test compound should be compared with results in control animals, which are not treated with the test compound. Administration of candidate IFT function restorative agents to the mammal can be carried out as described below, for example.

IFT function restorative agents can be identified with high throughput assays to detect IFT activity, e.g., the ability to grow cilia or flagella. For example, this restoration can be effected by small molecules binding directly to the IFT particle protein, or by binding of small molecules to other essential polypeptides in a biochemical pathway in which IFT participates.

The invention also provides methods of identifying agents (such as compounds, other substances, or compositions) that affect, or selectively affect, (such as inhibit, restore or otherwise modify) the activity of and/or expression of IFT particle polypeptides (or the IFT particle itself), by contacting an IFT polypeptide or the nucleotide sequence encoding the same with the agent and then measuring the activity of IFT, e.g., intraflagellar transport. In a related aspect, the invention features a method of identifying agents (such as compounds, other substances or compositions comprising same) that affect (such as inhibit, restore, or otherwise modify) the activity of and/or expression of nucleic acids encoding IFT particle polypeptides, by measuring the activity of and/or expression of IFT in the presence of the agent or after the addition of the agent in: (a) a cell line into which has been incorporated a recombinant construct including the nucleotide sequence of the nucleic acid encoding an IFT polypeptide or an allelic variation thereof; or (b) a cell population or cell line that naturally selectively expresses IFT polypeptides, and then measuring the activity of IFT polypeptides (or intraflagellar transport as a whole) and/or the expression thereof.

Since the IFT nucleic acids described herein have been identified, they can be cloned into various host cells (e.g., fungi, E. coli, or yeast) for carrying out such assays in whole cells.

To identify compounds that modulate expression of the IFT gene the test compound(s) can be added at varying concentrations to the culture medium of cells. Such test compounds can include small molecules (typically, non-protein, non-polysaccharide chemical entities),

polypeptides, and nucleic acids. The expression of IFT nucleic acid is then measured, for example, by Northern blot PCR analysis or RNAse protection analyses using a nucleic acid molecule of the invention as a probe. The level of expression in the presence of the test molecule, compared with the level of expression in its absence, will indicate whether or not the test molecule alters the expression of IFT. Because IFT is essential for survival of many organisms (e.g., mammals), test compounds that inhibit the expression and/or function of IFT nucleic acids are expected inhibit growth of, or kill, the cells that express IFT.

Some specific embodiments of the present invention relate to assay methods for the identification of anti-IFT agents using assays for anti-IFT agents which may be carried out both in whole cell preparations and in ex vivo cell-free systems. In each instance, the assay target is the IFT nucleotide sequence and/or the IFT polypeptide. Test compounds which are found to inhibit the IFT nucleotide sequence and/or IFT polypeptide in any assay method of the present invention are thus identified as potential or candidate anti-IFT agents. It is expected that the assay methods of the present invention will be suitable for both small and large-scale screening of test compounds as well as in quantitative assays such as serial dilution studies where the target IFT nucleotide sequence or the IFT polypeptide are exposed to a range of test compound concentrations.

A variety of protocols for detecting and measuring the expression of IFT, using either polyclonal or monoclonal antibodies specific for the protein, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on IFT polypeptides is suitable; alternatively, a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R et al. (1990, Serological Methods, A Laboratory Manual, APS Press, St Paul, MN) and Maddox, DE et al. (1983, J. Exp. Med. 158:121).

Pharmaceutical Formulations

The present invention also provides a pharmaceutical composition for treating an individual in need of such treatment of a disease caused by abnormal functioning of at least one IFT protein (or that can be treated by inhibiting abnormal IFT protein activity); the treatment

method entails administering a therapeutically effective amount of an agent that affects (such as inhibits) the activity and a pharmaceutically acceptable carrier, diluent, excipient, or adjuvant.

The pharmaceutical compositions can be used for humans or animals and will typically include any one or more of a pharmaceutically acceptable diluent, carrier, excipient, or adjuvant. The choice of pharmaceutical carrier, excipient, and diluent can be selected with regard to the intended route of administration and standard pharmaceutical practice. The pharmaceutical compositions can include as (or in addition to) the carrier, excipient, or diluent, any suitable binder(s), lubricant(s), suspending agent(s), coating agent(s), or solubilizing agent(s).

The invention includes pharmaceutical formulations that include a pharmaceutically acceptable excipient and an anti-IFT agent identified using the methods described herein. In particular, the invention includes pharmaceutical formulations that restore normal function to at least one IFT protein. Such pharmaceutical formulations can be used in a method of treating abnormal function of at least one IFT protein in an organism. Such a method entails administering to the organism a therapeutically effective amount of the pharmaceutical formulation, i.e., an amount sufficient to ameliorate signs and/or symptoms of abnormal IFT protein function. In particular, such pharmaceutical formulations can be used to treat abnormal IFT protein function in mammals such as humans and domesticated mammals (e.g., cows, pigs, dogs, and cats), and in plants. The efficacy of such IFT protein function restoration agents in humans can be estimated in an animal model system well known to those of skill in the art (e.g., mouse systems of abnormal IFT protein function).

Treatment includes administering a pharmaceutically effective amount of a composition containing IFT protein function restoration agent to a subject in need of such treatment, thereby restoring normal IFT protein function in the subject. Such a composition typically contains from about 0.1 to 90% by weight (such as 1 to 20% or 1 to 10%) of a IFT protein function restoration agent of the invention in a pharmaceutically acceptable carrier.

Solid formulations of the compositions for oral administration can contain suitable carriers or excipients, such as corn starch, gelatin, lactose, acacia, sucrose, microcrystalline cellulose, kaolin, mannitol, dicalcium phosphate, calcium carbonate, sodium chloride, or alginic acid. Disintegrators that can be used include, without limitation, micro-crystalline cellulose, corn starch, sodium starch glycolate and alginic acid. Tablet binders that can be used include acacia, methylcellulose, sodium carboxymethylcellulose, polyvinylpyrrolidone (Povidone®),

hydroxypropyl methylcellulose, sucrose, starch, and ethylcellulose. Lubricants that can be used include magnesium stearates, stearic acid, silicone fluid, talc, waxes, oils, and colloidal silica.

Liquid formulations of the compositions for oral administration prepared in water or other aqueous vehicles can contain various suspending agents such as methylcellulose, alginates, tragacanth, pectin, kelgin, carageenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also include solutions, emulsions, syrups and elixirs containing, together with the active compound(s), wetting agents, sweeteners, and coloring and flavoring agents. Various liquid and powder formulations can be prepared by conventional methods for inhalation into the lungs of the mammal to be treated.

Injectable formulations of the compositions can contain various carriers such as vegetable oils, dimethylacetamide, dimethylformamide, ethyl lactate, ethyl carbonate, isopropyl myristate, ethanol, polyols (glycerol, propylene glycol, liquid polyethylene glycol, and the like). For intravenous injections, water soluble versions of the compounds can be administered by the drip method, whereby a pharmaceutical formulation containing the anti-IFT agent and a physiologically acceptable excipient is infused. Physiologically acceptable excipients can include, for example, 5% dextrose, 0.9% saline, Ringer's solution, or other suitable excipients. For intramuscular preparations, a sterile formulation of a suitable soluble salt form of the compounds can be dissolved and administered in a pharmaceutical excipient such as Water-for-Injection, 0.9% saline, or 5% glucose solution. A suitable insoluble form of the compound can be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, such as an ester of a long chain fatty acid, (e.g., ethyl oleate).

A topical semi-solid ointment formulation typically contains a concentration of the active ingredient from about 1 to 20%, e.g., 5 to 10% in a carrier such as a pharmaceutical cream base. Various formulations for topical use include drops, tinctures, lotions, creams, solutions, and ointments containing the active ingredient and various supports and vehicles.

The optimal percentage of the IFT protein function restoration agents in each pharmaceutical formulation varies according to the formulation itself and the therapeutic effect desired in the specific pathologies and correlated therapeutic regimens. Appropriate dosages of the IFT protein function restoration agents can be determined by those of ordinary skill in the art of medicine by monitoring the mammal for signs of disease amelioration or inhibition, and increasing or decreasing the dosage and/or frequency of treatment as desired. The optimal

amount of the IFT protein function restoration agent used for treatment of conditions caused by or contributed to by abnormal IFT protein function depends upon the manner of administration, the age and the body weight of the subject, and the condition of the subject to be treated. Generally, the IFT protein function restoration compound is administered at a dosage of 1 to 100 mg/kg body weight, and typically at a dosage of 1 to 10 mg/kg body weight.

B. Diagnostic Uses

Diagnosis of Defective or Absent IFT Particle Proteins

PCR to Detect Missing or Defective IFT Particle Protein Genes

RNA is extracted from patient biopsy material, e.g. nasal scrapings, and reverse transcribed into cDNA. PCR reactions are carried out on this cDNA using pairs of PCR primers designed from the sequence of the human homologue of each of the IFT particle protein genes. If the patient lacked mRNA for any of the genes, then no product would be amplified. The PCR products can also be analyzed by sequencing or other standard methods to identify nucleotide changes. The Tg737 mutant mouse has reduced levels of the mRNA derived from the mouse Tg737/IFT88 gene (Moyer et al., 1994).

Alternatively, PCR can be used to amplify exons of genes encoding IFT particle proteins from genomic DNA purified from patient blood. These amplified products can then be sequenced or analyzed by other standard methods to identify nucleotide changes.

Antibody Assays to Identify Missing IFT Proteins

Protein extracts would be prepared from patient biopsy material, e.g. nasal scrapings, and analyzed by enzyme-linked immunosorbent assay (ELISA) or by western blot analysis using antibodies to the IFT particle proteins. The amount of protein in the patient sample would be compared to normal controls to determine if the patient lacks a particular IFT particle protein. Chlamydomonas cells with a mutation in the IFT88 gene lack the IFT88 protein and IFT57 protein (Pazour et al., 2000) and the Tg737 mutant mouse has reduced amounts of full length Tg737/IFT88 protein (Pazour unpublished; Taulman et al., 2001).

C. Use in Agriculture

Combating phytopathogenic nematodes

Phytopathogenic nematodes are responsible for tens of billions of dollars lost each year for farmers throughout the world (Williamson, *Curr Opin Plant Biol*, 2:327-331, 1999). A new way to combat these parasites is the use of internalizable agents that interfere with the nematode's ability to sense its environment by disrupting the structure and function of nematode sensory cilia. Nematodes largely sense their environment by chemosensation, a process that occurs at the sensory cilia of sensory neurons (Troemel, *Bioessays*, 21:1011-1020, 1999). Loss of function of these cilia will have multiple effects, including the loss of ability to locate the host plant as well as reproductive mates.

Small Molecule Inhibitors of Intraflagellar Transport

Intraflagellar Transport (IFT) is necessary for both the construction and continued maintenance of cilia and flagella in diverse organisms, including nematodes (Rosenbaum and Witman, unpublished manuscript, 2001). If IFT is interrupted, cilia shorten and become dysfunctional. We propose to block IFT with one or more small molecules. Large numbers of small molecules will be screened for their ability to disrupt IFT and thus, sensory cilia function. The IFT motors, kinesin-II and cytoplasmic dynein 1b would be likely targets but other IFT machinery might also be involved. There is already precedence for this strategy. Monastrol is a small organic molecule that was found via screening to selectively inhibit a subgroup of kinesins known as Eg5 or bimC (Mayer et al, *Science*, 286:971-974, 1999). Eg5 inhibition by monastrol is specific and does not affect the behavior of other kinesins.

The screening process would identify the loss of sensory cilia function. Normal nematode sensory cilia are exposed to the environment and have the unique ability to take up membrane-soluble fluorescent dyes such as DiO and DiI (Molecular Probes, Inc). The cilia-dependent uptake of these dyes into the sensory neurons allows for easy and fast screening using fluorescence microscopy and has been used to identify a relatively large number of sensory cilia mutants (Starich et al, *Genetics*, 139:171-188, 1995). In mutants where the sensory cilia are structurally defective and fail to extend out into the environment, no dye is taken up into the sensory neurons. This same screen can be used to identify agents that cause normal, full-length sensory cilia to shorten to the point at which they are no longer exposed to the environment. As

for the anti-IFT agents, many possible small molecules can be screened. Since monastrol binds and specifically inhibits Eg5 kinesin, it would be logical to try variations of this molecule since slight changes in chemical structure can result in a molecule that binds and inhibits kinesin-II. Molecules that affect nematode IFT would likely be added topically to the plant or the soil. Another, more specific approach involves identification of small peptide(s) that block IFT.

Identification of one or more small peptides that block nematode IFT is particularly attractive because host plants can be transformed with the appropriate vector so that the host plant makes the inhibitory peptide(s). Transformation would be mediated by *Agrobacterium tumefaciens* with a standard vector such as pCAMBIA1380 or pCAMBIA1390 (Center for the Application of Molecular Biology to International Agriculture) encoding the inhibitory peptide. Different promoters can be chosen which would result in either constitutive expression or tissue-specific expression of the peptide gene. For example, many nematodes attack their host plant at the roots. Root-specific expression, therefore, is desirable.

RNA Interference of Nematode IFT

Known as RNA interference or RNAi, the introduction of gene-specific sequences of double-stranded RNA (dsRNA) result in a specific gene silencing event which can block the ability of that specific gene to generate protein. In the nematode, this specific dsRNA can be introduced simply by ingestion (Timmons and Fire, *Nature*, 395:854, 1998; Timmons, Court and Fire, *Gene*, 263:103-112, 2001). IFT genes are specifically targeted in the nematode by transforming plant hosts so that they produce dsRNA derived from portions of genes encoding nematode IFT machinery, including kinesin-II and dynein 1b subunits as well as the IFT raft proteins. This strategy allows for very specific targeting of nematodes that attack and ingest part of the host plant. DNA/RNA sequences of portions of the IFT machinery will be chosen which do not have high homology to the homologous genes in other classes of organisms so as to protect various animals and humans from potential RNAi. Indeed, noncoding introns may be particularly useful for this purpose.

Generating dsRNA via plant transformation with a stable vector has already been achieved in Arabidopsis (Chuang and Meyerowitz, *Proc Natl Acad Sci*, 97:4985-4990, 2000). For transformation, suitable vectors such as pCAMBIA vectors can be used. To generate the dsRNA, as little as 100 base pairs of coding or noncoding sequence would be linked to double-

stranded complementary sequence so that when the RNA is generated, the two matching RNA sequences would base pair with one another and form double stranded RNA. Attacking nematodes would ingest this dsRNA which would lead to RNA silencing of IFT genes which would result in the loss of the sensory cilia. We would likely want to use a strong promoter so that the concentration of dsRNA in the plant cells would be relatively high. The actual dose needed to affect loss of ciliary function would be measured in a laboratory setting. One positive aspect of RNAi technology is that small amounts of RNA appear to be amplified within the nematode and transmitted adjacently (Fire et al., *Nature*, 391:806-811, 1998).

IFT and Insect Pests

All insect pests utilize sensory cilia for important life functions, including location of mates and identification of food sources. Sensory cilia (ciliated sensory neurons) are the organelles for smell, taste and hearing in insects. A compound that inhibited IFT would block sensory cilia function and/or assembly, leaving an insect unable to smell, unable to taste, and unable to hear. Insects deprived of these senses would be unable to locate mates; as a result, their populations would be controlled. Many insects deprived of these senses also would be unable to sense food, and thus would not eat, or would eat plants that would be harmful to them; this would both reduce insect populations and reduce damage to crop plants. Therefore, a compound that blocked IFT would be a very effective pesticide. Such a compound can be applied by spraying of fields, plants, insect habitats, etc., or by any other preferred method.

For example, the male cabbage looper moth (*Trichoplusia ni*) uses ciliated sensory neurons, called sensilla, which are located on its antennae, to find a female moth (Borroni and O'Connell, *J. Comp. Physiol.*,170:691-700, 1992). The female moth releases a specific pheromone, which the male then detects by means of its sensilla and follows upwind until it locates the female, with which it then mates. By disrupting the assembly and/or function of the male's sensilla, an anti-IFT compound would render the male moth unable to find the female, and the thus control the population of this major pest.

In another example, mosquitoes use olfactory receptors located on sensory cilia to detect CO_2 and other attractants, and thus find their prey (Grant et al., *J. Comp. Physiol.*,177:389-396, 1995). Spraying mosquitoes with an anti-IFT compound, which would block sensory cilia assembly/function, would render the insect unable to locate its prey and obtain a blood meal.

Combating Insects

The approaches described above regarding nematodes can be adapted to combat insects. Insects that attack and consume plants are good targets for plant transformants that generate either anti-IFT peptides or double stranded RNA that is derived from insect genes encoding IFT proteins. Blocking IFT in the insect should result in at least two desirable effects. First, the chordotonal organ, also known as Johnston's organ, contains a ciliated structure that is necessary for sound transduction. Second, insect sperm is flagellated and motile. Interruption of IFT should result in (1) deaf insects, which will affect communication with other insects, including potential mates and (2) immotile sperm, which will the males of the species infertile. Both of these effects can dramatically reduce the population of an insect species. Additional cells/tissues that can be affected by blocking IFT include the mechanosensory bristles. Many mutations that affect sound transduction by the chordotonal organ also affect touch transduction by these tactile bristles (Eberl, Hardy, and Kernan, *J. Neurosci.*, 20:5981-5988, 2000); these results suggest that IFT, or portions of the IFT machinery, is involved with touch transduction in insects. Losing the sense of touch as a result of anti-IFT agent(s) would result in an obvious disadvantage for the organism.

Examples

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

Example 1. Identification of the Genes Encoding Chlamydomonas IFT Particle Proteins and Homologous Proteins in Other Species

Purification and Microsequencing of Chlamydomonas IFT Particle Proteins

16S IFT particles were purified from Chlamydomonas flagella as described in Cole et al. (1998), further purified by two-dimensional gel electrophoresis and transferred to ImmobilonP^{SQ} (Millipore Corp., Bedford, MA). The spots corresponding to each of the IFT particle proteins were excised and digested with trypsin. Tryptic peptides were eluted from the membrane and fractionated by high performance liquid chromatography. Pure peptides, identified by mass

spectrometry, were subjected to microsequence analysis in the UMMS Protein Sequencing Facility.

Cloning Genes Encoding *Chlamydomonas* IFT Particle Proteins

The peptide sequences from the purified IFT particle proteins were used to design PCR primers that were then used to amplify portions of the genes from *Chlamydomonas* genomic DNA or cDNA. These fragments of DNA were then used to obtain the rest of the genes by screening cDNA libraries by DNA hybridization, or the fragments sequenced and the sequences used to design additional primers that were used in 5' and 3' RACE (rapid amplification of cDNA ends) reactions to PCR amplify the remainder of the genes.

Obtaining Mouse IFT57 sequence

The predicted peptide from *Chlamydomonas* IFT57 was used to search the ESTS portion of Genbank. One positive ESTS clone (Accession # AA763980) was purchased from ATCC and sequenced.

Identifying Homologous Genes

The predicted peptide sequences of the *Chlamydomonas* IFT particle proteins were used to search the protein and ESTS portions of Genbank with BLAST to identify homologous proteins. Typically human and *Chlamydomonas* homologues were 30-50% identical at the protein sequence level. If a human homologue was not present in either the protein or ESTS portion of Genbank, then the *Chlamydomonas* sequences (and any mammalian homologues found above) were used to search the draft human genome sequence to identify homologous loci. The proteins predicted to be encoded at these loci were compiled using the homology to the *Chlamydomonas* (or non-human mammalian) homologue as a guide. Similar methods were used to identify *Caenorhabditis elegans* homologues and can also be used to identify genes encoding IFT particle proteins in any organism (e.g. *Giardia*, *Plasmodium*, *Drosophilia*) for which extensive DNA sequence data has been obtained.

Example 2. Correction of IFT Defects in Retinal Cells by Gene Therapy

Defects in IFT particle proteins cause retinal degeneration in mice (Pazour et al., 2001) and are likely to be a cause of degenerative retinal disease leading to blindness in humans (Rosenbaum and Witman, 2001). In such cases vision may be corrected by gene therapy methods in which a vector containing the wild-type gene for the defective IFT particle protein is injected subretinally and the IFT particle protein gene thereby transfected into the photoreceptor cells and expressed.

The vector can be any non-viral or viral vector, e.g., recombinant adeno-associated virus (rAAV) vector (Ackland et al., 2001), which has the ability to transfect the target retinal cells. Expression of the IFT gene would be driven by its endogenous promoter, or by a photoreceptor cell-specific promoter, e.g., the opsin promoter, or by another promoter, such as the immediate-early cytomegalovirus (CMV) enhancer-promoter, which would be placed upstream of the IFT gene in the construct. Plasmid DNA containing this construct would be packaged into the vector.

In the case of rAAV, this can be achieved by inserting the gene into a basic rAAV vector plasmid, e.g., pTR-UF5, and transfecting the plasmid into a suitable human host cell in culture, e.g., human 293 cells, together with helper plasmid required for productive AAV infection and packaging of the rAAV DNA. After harvesting cells, the virus is extracted by a standard procedure, e.g., by freezing and thawing the cells, and then purified by iodixanol density gradient centrifugation followed by heparin-Sepharose agarose column chromatography (Hauswirth et al., 2000 Meth. Enzymol. 316: 743-761). Alternatively, any other procedure which results in efficient recovery of high quality virus may be used.

The purified vector would then be injected into the subretinal space underlying the central retina by means of an anterior chamber cannula inserted through a sclerotomy and monitored by microscopy (Bennett et al., 1999). Related protocols have been demonstrated to give stable, long-term expression of the transgene in non-human primate photoreceptor cells (Bennett, J. et al., 1999. PNAS 96: 9920-9925), without toxicity, and to restore vision in a large animal model – the dog -- in which retinal degeneration occurs as a result of a defect in the RPE65 gene (Acland et al., 2001).

Example 3. Correction of IFT Defects in Airway Epithelial Cells by Gene Therapy

Because IFT particle proteins are necessary for the formation and maintenance of all cilia and flagella, defects in IFT particle proteins are expected to prevent normal assembly and functioning of respiratory tract cilia. Defects in assembly and functioning of respiratory tract cilia lead to human disorders such as primary ciliary dyskinesia (Afzelius and Mossberg, 1995), which are characterized by bronchiectasis, chronic bronchitis, and chronic sinusitis. Respiratory tract cilia arise from epithelial cells, which are exposed on the surface of the airway. Thus, these cells should be ideal targets for gene therapy to correct pulmonary disease due to a defect in an IFT particle protein. In this case, viral or nonviral vectors containing the IFT gene can be delivered topically to the airways via direct liquid instillation (Yonemitsu et al., 2000) or via an aerosol, such as can be produced by a nebulizer (Gautam et al., 2001).

The wild-type IFT particle protein gene to be transfected into the target cells can be contained in a plasmid ("IFT gene plasmid") and can be under the control of its endogenous promoter or another promoter, e.g., the human CMV early promoter/enhancer element. In one approach, the purified plasmid can be inserted into a rAAV vector plasmid for production of rAAV in a suitable human host cell, followed by purification of the rAAV as described above for gene therapy methods to correct a defect in the retina. The purified rAAV can then be delivered to the airway epithelial cells by instillation of a solution containing the rAAV, or by aspiration of an aerosol containing the rAAV. Related approaches might use recombinant lentiviral vectors (Kobinger et al., 2001) or recombinant Sendai virus vectors (Yonemitsu et al., 2000) to transfer the IFT gene to the airway epithelial cells. Both of these vectors have been shown to efficiently transduce airway epithelial cells (Kobinger et al., 2001; Yonemitsu et al., 2000).

Alternatively, the IFT gene plasmid can be complexed with other components to create a nonviral vector that can fuse with the airway epithelial cell membrane and deliver the plasmid DNA to the cytosol, from where it would then move to the nucleus. For example, the IFT gene plasmid DNA can be mixed with cationic lipids to produce a DNA-cationic lipid complex or "liposome" that can be delivered as an aerosol or via liquid instillation. Additional components, such as protamine or proteins, may be added to increase efficiency of delivery of the vector to the target cell (e.g., Sorgi, FL et al., 1997. Gene Therapy 4:961-968). Gene transfer efficiency is 10-200 times higher using liposomes containing specific peptides – e.g., an integrin-binding motif – than using liposomes alone (Scott et al., 2001. J. Gene Med. 3:125-134). In another approach, the DNA can be complexed with polyethylenimine (PEI) to create a PEI-DNA

complex. PEI-DNA complexes have been demonstrated to be effective for the aerosol delivery of reporter genes to the lungs of mice, and may be more effective than liposomes (A. Gautam et al., 2000. Mol. Therapy 3:551-556).

With either viral or nonviral vectors, the efficiency of delivery of the DNA to the target cells may be enhanced by directing the vector to a specific target, e.g., the extracellular ATP/UTP receptor, termed P2Y2-R, which internalizes into the cell via clathrin-coated pits upon agonist stimulation (Boucher, RC. 1999. J. Clin Invest. 103:441-445).

Example 4. Determining the Role of IFT88 For Assembly of Cilia and Flagella

Introduction

Defects in the Tg737 gene cause kidney and liver defects in mice that are very similar to those seen in humans with autosomal recessive polycystic kidney disease (ARPKD) (Moyer et al., 1994). This disease affects ~ 1 in 10,000 children born each year and may be responsible for a much higher proportion of stillbirths and prenatal deaths (Blyth and Ockenden, 1971; Cole et al., 1987). The function of the Tg737 protein is unknown. Here we identify a protein in *Chlamydomonas* that is homologous to Tg737 and show that it is required for assembly of flagella.

The epithelial cells lining the collecting tubules of the kidney have very well developed primary cilia (Andrews and Porter, 1974). The role of these cilia is unknown; however they extend into the lumen of the tubule and may serve as sensory appendages. Precedence for primary cilia serving a sensory role is well established in vision and olfaction, as the outer segments of the rod and cone cells of the eye and the olfactory cilia of the nose have evolved from cilia and have retained primary cilia characteristics, e.g. the 9+0 microtubule arrangement. Primary cilia in other organisms such as *Caenorhabditis elegans* also serve a sensory role (Perkins et al., *Dev. Biol.*, 117:456-487, 1986; White et. al., *Phil. Trans. R. Soc. Lond.*, 275:327-348, 1976).

Eukaryotic cilia and flagella are built and maintained by a process called intraflagellar transport (IFT) (Rosenbaum et al., 1999). Most well characterized in *Chlamydomonas*, IFT is a rapid movement of particles along the axonemal microtubules of cilia and flagella. The outward movement of these particles from the cell body to the tip of the flagellum is driven by FLA10

kinesin-II (Kozminski et al., 1993; 1995), whereas the transport of the particles from the tip back to the cell body is driven by DHC1b/DHC2 cytoplasmic dynein (Pazour et al., 1998, 1999; Porter et al., 1999). The particles that are transported by IFT are composed of at least 17 protein subunits (Piperno et al., 1997; Cole et al., 1998). The functions of the individual subunits are not known but the proteins are conserved between green algae, nematodes, and vertebrates (Cole et al., 1998; Rosenbaum et al. 1999). In this manuscript, we describe the cloning of the IFT88 subunit of the *Chlamydomonas* IFT particle and show that cells missing this gene do not assemble flagella. We further show that IFT88 is homologous to the polycystic kidney disease gene Tg737 and that mice with mutations in this gene have shorter than normal primary cilia in their kidney.

Materials And Methods

Purification and Microsequencing of *Chlamydomonas* IFT88

16S IFT particles were purified from *Chlamydomonas* flagella as described in Cole et al. (1998). The IFT88 subunit was further purified by two-dimensional gel electrophoresis and transferred to ImmobilonP^{SQ} (Millipore Corp., Bedford, MA) as described previously (Cole et al., 1998). The spot corresponding to IFT88 was excised and digested with trypsin. Tryptic peptides were eluted from the membrane and fractionated by high performance liquid chromatography. Pure peptides, identified by mass spectrometry, were subjected to microsequence analysis in the UMMS Protein Sequencing Facility.

Cloning IFT88

Portions of the IFT88 peptide sequence (LEGETDQA and GIDPYCVE) were used to design two degenerate oligonucleotide PCR primers (GA[A/G] AC[C/G/T] GA[C/T] CA[A/G] GC[C/G/T] GA[C/T] AA[A/G] TA and GC [C/T]TC [A/C/G]AC [A/G]CA [A/G]TA [A/C/G]GG [A/G]TC [A/G]AT). These primers amplified a 365-bp fragment of genomic DNA that contained parts of two exons and a 132-bp intron. This fragment of genomic DNA was used to screen a *Chlamydomonas* cDNA library made from cells undergoing division (Pazour and Witman, unpublished). Two positive clones were identified and sequenced by primer walking. These two clones were similar except for the sequences at their 5'-ends. IFT88cDNA-1 was longer than IFT88cDNA-2 and appeared to have a short region of poly-A inappropriately fused

to the 5'-end, probably the result of a cloning artifact. One *Chlamydomonas* IFT88 EST clone is in Genbank (accession number AV395576). This EST sequence, which is from the 5' end of the gene and overlaps the cDNA clones, was used to define the 5'-end of the cDNA sequence.

Four independent BAC clones (40-B3, 11-O21, 24-F2, and 27-M3) were found in the Genome Systems (St Louis, MI, USA) *Chlamydomonas* BAC library by Southern hybridization using the 365-bp fragment of *IFT*88 genomic DNA as a probe. These four BAC clones were purchased from Genome Systems. The presence of the IFT88 gene in the clones was confirmed by Southern blotting.

Identification and Rescue of an IFT88 Mutant

DNA from each of the ~400 transformants in our insertional mutant collection (Pazour et al., 1995; 1998; Pazour and Witman, 2000) was cut with PstI and analyzed by Southern blotting with the 365-bp fragment *IFT*88 genomic DNA fragment. This probe detected an ~2.5-kb band in wild-type cells and all of the mutants except strain V79.

The motility/flagellar defect in V79 was rescued by transforming with BAC clones carrying the *IFT*88 gene. Transformation was performed by the glass bead method of Kindle (1990), and rescued cell lines were identified by restoration of their ability to swim. One rescued cell line was crossed to wild-type CC124 cells. Tetrads from this cross were dissected and analyzed by standard procedures (Levine and Ebersold, 1960; Harris, 1989) as described in Pazour et al. (1998). The flagellar phenotype was scored by light microscopy when the cells were in the early log phase of growth.

Electron Microscopy

Chlamydomonas cells were fixed in glutaraldehyde for EM (Hoops and Witman, 1983) and processed as described in Wilkerson et al. (1995). Tissues of anesthetized mice were fixed *in situ* by brief cardiac perfusion with 2.5% gluteraldehyde in 100 mM cacodylate buffer. The kidneys were removed and a small amount of additional fixative was injected under the capsule of the kidney. The kidneys were placed in additional fixative for 1 hour. At that time, the kidneys were sliced in half and further fixed for 2 days. The tissue was freeze fractured and metal impregnated as described in McManus et al. (1993).

Western Blotting

Whole cell extracts of wild-type and mutant cells were made by resuspending log-phase cells in SDS-sample buffer, heating at 50° C for 10 min, and repeatedly drawing the sample through a 26-guage needle to shear the DNA. Proteins were separated by SDS-PAGE, blotted onto polyvinylidene difluoride membranes, and probed with antibodies as described in Pazour et al. (1998). Antibodies used included mAb57.1, mAb81.1, mAb139.1, and mAb172.1, which are monoclonal antibodies against IFT particle proteins (Cole et al., 1998); FLA10N, which is specific for a kinesin-II motor subunit (Cole et al., 1998); DHC1b, which is specific for the heavy chain of DHC1b/DHC2 cytoplasmic dynein; and B-5-1-2, which is specific for alpha tubulin (Piperno et al., 1985).

Chlamydomonas Culture

Chlamydomonas strains used in this work included: g1 (*nit*1, *agg*1, *mt*+) (Pazour et al., 1995), CC124 (*nit*1, *nit*2, *mt*-), and CC1390 (*fla*2, mt-). The latter two strains are available from the *Chlamydomonas* Genetics Center (Duke University, Durham, NC). Strains generated in the course of this work included: V79 (*ift*88-1::*NIT*1, *mt*+), which was generated by random insertional mutagenesis of g1, V79/40-B3#2.5 (*ift*88-1::*NIT*1, *IFT*88, *mt*+) generated by transformation of V79 with BAC clone 40-B3, and 3276.2 (*ift*88-1::*NIT*1) which is a progeny from a cross between V79/40-B3#2.5 and CC124.

Mouse Genotyping

DNA was purified by digesting mouse tails with proteinase K, extracting once with 50% phenol / 50% chloroform, once with chloroform and then precipitating the DNA with ethanol. The genomic DNA was amplified using the RW450, RW451, and RW452 primer set described by Yoder et al. (1997). These primers amplified a 270-bp fragment from the wild-type locus and a 340-bp fragment from the mutant locus.

Digital Image Processing

Western and Southern blots were scanned from negative x-ray film with a Linotype-Hell Saphir Ultra 2 flatbed scanner and brought into Photoshop for cropping and contrast adjustment.

Scanning EM micrographs were scanned from positive Polaroid film in the same way. Transmission EM negatives were scanned with a Polaroid Sprint Scan 45 and brought into Photoshop for cropping, contrast adjustment and inversion from a negative to a positive image.

Results

Cloning and Sequencing of *Chlamydomonas* IFT88

In order to learn more about the structure and function of the proteins that make up the IFT particle, we cloned and sequenced the IFT88 protein, formerly known as p88 (Cole et al., 1998). To do this, *Chlamydomonas* IFT particles were purified from the matrix of isolated flagella by sucrose density gradient centrifugation and two-dimensional gel electrophoresis. IFT88 was cleaved by trypsin and two internal peptides were microsequenced (Cole et al., 1998), yielding the sequences AATNLAFLYFLEGETDQADKYSEMALK and SLFNEAAGIDPYCVEAIYNLGLVSQR. Degenerate PCR primers were designed from these sequences and used to amplify a fragment of genomic DNA. A cDNA library was screened with the genomic fragment and the resulting clones were sequenced by primer walking. Southern blots indicated that there is only one copy of the *IFT*88 gene in the *Chlamydomonas* genome.

Sequence analysis showed that the *IFT*88 cDNA contains a 2346-nt open reading frame that is predicted to encode an 86.3-kD protein with a pI of 5.87. Perfect matches to both IFT88 tryptic peptides are found in the open reading frame of this cDNA, rigorously confirming that these clones encode the *Chlamydomonas* IFT88 protein. No discernable motifs were identified within the sequence except for the presence of 10 tetratricopeptide repeats (TPR). TPRs are degenerate 34-amino acid repeats (Lamb et al., 1995), present in tandem arrays of 3-16 units that are predicted to form amphipathic helices (Hirano et al., 1990). The first three TPR motifs are found closely spaced between amino-acid residues 185-294. The other seven TPR motifs occur without spacing between amino-acid residues 441-676.

Chlamydomonas IFT88 Is Homologous to a Mouse Polycystic Kidney Disease Gene

Blast searches with the *Chlamydomonas* IFT88 protein sequence indicate that it is very similar to the mouse (41% identical; BLAST E=e-148) and human Tg737 (40% identical; BLAST E=e-146) proteins. Mice with defects in Tg737 have severe polycystic kidney disease and die within a few weeks of birth. The protein also is homologous to proteins predicted by

ESTs from zebra fish and swine and fragments of preliminary *C. elegans* and *D. melanogaster* genomic sequence. IFT88 and Tg737 are likely to be functionally equivalent orthologues as the similarity between the *Chlamydomonas* and mammalian proteins is robust and distributed over the entire coding region and not just within the TPR domains. 40% identity is very typical of the amount of similarity seen between other *Chlamydomonas* and mammalian orthologues that encode flagellar proteins (Pazour, Dickert and Witman, manuscript in preparation).

IFT88 Is Required for Flagellar Assembly

To learn more about the function of IFT88 in cells, we searched our collection of *Chlamydomonas* insertional mutants (Pazour et al., 1995, 1998; Pazour and Witman, 2000) for a cell line with a defect in this gene. The insertional mutants were made by transforming cells with DNA carrying a selectable marker. In *Chlamydomonas*, transforming DNA is integrated randomly throughout the genome and disrupts genes at the site of integration. DNA was isolated from ~400 insertional mutants having behavioral or motility defects, and was screened by Southern blotting using a fragment of *IFT*88 genomic DNA as probe. One cell line (V79) was identified that had an insertion in the *IFT*88 gene. The fact that the single hybridizing band in wild-type cells was split into two bands in the mutant indicated that the selectable marker integrated into the gene within the region covered by the probe and did not result in a large deletion of the genome at the site of integration. The mutant allele was termed *ift*88-1.

The *ift*88-1 cells grew at the same rate as wild-type cells, indicating that IFT88 is not required for processes essential for growth or cell division. However, in contrast to wild-type cells that normally have two ~10-μm flagella extending from the anterior end of the cell body, the *ift*88-1 cells completely lack flagella. Electron microscopic analysis of these cells showed that the basal bodies were structurally normal but the flagella did not extend beyond the transition zone. In some cells, the membrane covering the flagellar tips was tightly apposed to the microtubules with no material between them and the membrane. In other cells, the flagellar stubs were slightly swollen and contained fragments of microtubules in random orientations. However, in contrast to the IFT mutants *fla*14 (Pazour et al., 1998) and *dhc*1b (Pazour et al., 1999), no accumulation of IFT particles was observed in any of the flagellar stubs.

To determine the effect of the lack of IFT88 on the IFT particle and the IFT motors, we examined whole cell extracts by western blotting. The IFT particle is made up of two large complexes. Complex A is composed of four to five proteins and includes IFT139. Complex B is

composed of IFT88 and ten other proteins including IFT172, IFT81 and IFT57. The complex A protein IFT139 is enriched in the mutant suggesting that the gene may be upregulated in the mutant cells. The mutation has little or no effect on the levels of complex B proteins IFT172 and IFT81, but causes a significant decrease in IFT57, another complex B protein. The cellular levels of the IFT motors FLA10 kinesin-II and DHC1b are not affected by the *ift*88 mutation.

To be certain that the flagellar assembly defect is caused by the mutation in *IFT*88 and is not the result of another mutation elsewhere in the genome, we transformed the mutant cells with BAC clones carrying the *IFT*88 gene. Three independent BAC clones (40-B3, 24-F2, and 27-M3) complemented the flagellar defect. The complemented cell lines swam like wild-type cells and had IFT. One of the rescued cell lines was crossed to a wild-type cell line and 26 tetrads were dissected. All four products of one tetrad and a single product of the remaining 25 tetrads were analyzed by Southern blotting. Because the transformed copy of the *IFT*88 gene inserted at a site unlinked to the original locus, the inserted DNA segregated independently from the original gene, allowing offspring to carry zero, one, or two copies of the wild-type gene. Cells that carry at least one copy of wild-type *IFT*88 have normal flagella and motility, whereas those that carry no copies of wild-type *IFT*88 lack flagella and are non-motile. These data indicate that the flagellar defect is tightly linked to the *ift*88 mutation and is almost certainly the result of it.

Primary Cilia in the Kidney of Tg737 Mice Are Shorter Than Normal

Primary cilia are present on many cells in the mammalian body (Wheatley, 1995; Wheatley et al., 1996), and are particularly well developed in the kidney (Andrews and Porter, 1974). Inasmuch as *Chlamydomonas* IFT88 is necessary for assembly of flagella and is homologous to mammalian Tg737, and because a defect in mouse Tg737 leads to kidney disease, it was of great interest to determine if the defect in Tg737 affected formation of the primary cilia in the kidney. In wild-type rats, the cilia are ~2.5 μm long and are found in the proximal tubule, the loop of Henle, the distal tubules, and the collecting ducts (Andrews and Porter, 1974). In wild-type mice these cilia are less than 5 μm long and similarly distributed (Flood and Totland, 1977).

We obtained the hypomorphic Tg737-mutant mice from Oak Ridge National Laboratory and examined the kidneys of 4-day and 7-day old pups by scanning electron microscopy.

Numerous monociliated cells were observed in the kidneys of both wild-type (+/+) and homozygous mutant (-/-) mice, but the cilia in the mutant kidneys were much shorter. To quantify this difference, the cilia lengths were measured from scanning electron micrographs taken from the tubules distal to the proximal tubule. The proximal tubule was avoided because it contains a thick brush border that can obscure a micron or more of cilia length. The tubules distal to the proximal tubule have only sparse microvilli that do not obscure cilia, and the cilia in these regions are uniform in length (Andrews and Porter, 1974). These cilia in wild-type mice were 3.1+/- 1.4 μm and 3.5 +/- 1.7 μm long at 4 and 7 days respectively, whereas these cilia in the mutant mice were 1.0 +/- 0.6 μm and 1.3 +/- 0.6 μm at 4 and 7 days respectively. These values represent minimum lengths as it is difficult to accurately measure cilia that are lying at different angles in the tubules. However, the differences are quite large and are significant at the >99% level. Thus, Tg737, like its IFT88 homologue in *Chlamydomonas,* plays an essential role in assembly of the primary cilium in the mouse.

The IFT88 Gene Is Required for Flagellar Assembly in *Chlamydomonas*

Chlamydomonas cells lacking the *IFT*88 gene do not assemble flagella, indicating that the IFT88 protein is required for flagellar assembly. This is the first *Chlamydomonas* IFT particle subunit to be shown to be required for ciliary assembly. Loss of IFT88 causes a substantial decrease in the amount of IFT57 relative to other IFT particle proteins in the cytoplasm, suggesting that IFT88 is important for assembly of at least a portion of the IFT particle. Thus, IFT may be blocked at a very early stage in the *ift*88-1 mutant. Consistent with this, IFT particles do not accumulate in the flagellar stubs of the *ift*88 mutant, in sharp contrast to the dramatic accumulation of apparently intact IFT particles in the flagella of mutants with defects in cytoplasmic dynein DHC1b/DHC2 (Pazour et al, 1999) or the dynein light chain LC8 (Pazour et al., 1998). Alternatively, IFT88 may have a vital role in the attachment of the IFT particle to its cargo or to the anterograde IFT motor FLA10-kinesin-II; in either case, loss of IFT88 would preclude flagellar assembly. It is also possible that IFT88 is essential for transduction of a signal that is necessary for flagellar assembly.

Tg737, the Mouse IFT88 Homologue, Is Required for Assembly of the Primary Cilia in Kidney.

We have shown that IFT88 is highly similar to the mouse and human Tg737 proteins and that mice with defects in Tg737 have defective cilia in their kidneys. Tg737 was identified at

Oak Ridge National Laboratory by random insertional mutagenesis of mice. Hypomorphic mutations in Tg737 cause kidney disease and death within a few weeks of birth. The phenotype of this mutation closely resembles human ARPKD in that the mice develop cystic kidneys and have hepatic biliary disease which is also common in human patients with ARPKD (Moyer et al., 1994). The mice develop large cysts in their collecting ducts and are unable to concentrate urine (Yoder et al., 1996; 1997). Null alleles of Tg737 have a more severe phenotype and cause the mice to die during mid-gestation (Murcia et al., 2000). The phenotype caused by the null Tg737 mutation closely resembles the phenotype of kinesin-II knockout mice (Nonaka et al., 1998; Marszalek et al., 1999, Takeda et al., 1999). Both the kinesin-II and Tg737 null mice have left-right asymmetry defects, lack cilia on the embryonic node, and die during mid gestation. Our finding that IFT88 is required is required for ciliary assembly provides the first evidence that the lack of nodal cilia on embyros of Tg737 null mutant mice is a direct result of a defect in IFT.

Primary cilia are extremely widely dispersed throughout the mammalian body. The only cells that are known NOT to contain primary cilia are hepatocytes, and differentiated cells of myeloid or lymphoid origin (Wheatley, 1995; Wheatley et al., 1996). The primary cilia in kidney tubules and ducts (Andrews and Porter, 1974) and hepatic biliary ducts (Motta and Fumagalli, 1974) are unusually long and project into the lumens of these structures. The role of the primary cilia in the kidney or hepatic ducts is not known but has been suggested to be sensory (Roth et al., 1988). The most studied primary cilia are the outer segments of rod and cone photoreceptor cells and the olfactory cilia in the nasal cavity. In these examples, the role of the primary cilia is clearly to serve as an appendage to concentrate sensory machinery. *C. elegans* also makes extensive use of primary cilia to detect osmolarity gradients and chemical signals (White et. al., *Phil. Trans. R. Soc. Lond.*, 275:327-348, 1976; Lewis and Hodgkin, *Comp. Neur.*, 172:489-510, 1977). Primary cilia on other cells may similarly have a sensory role. Supporting this idea, the somatostatin 3 receptor has recently been localized to primary cilia in the brain (Handel et al., 1999). Kidney epithelial cells sense multiple extracellular signals including peptide hormones like angiotensin and ions like chloride (reviewed in Gunning et al., 1996). Whether any of these sensory receptors are localized to the primary cilia of the vertebrate kidney remains to be determined.

C. elegans homologues of the human polycystic kidney disease genes, PKD1 and PKD2, are localized to sensory cilia (Barr and Sternberg, 1999). Humans with mutations in PKD1 and

PKD2 develop kidney disease similar to that caused by Tg737 mutations in mice. PKD1 and PKD2 are transmembrane proteins that interact with each other (Qian et al., 1997; Tsiokas et al., 1997). PKD1 has a large extracellular domain that is thought to bind an unknown ligand (Hughes et al., 1995; The International Polycystic Kidney Disease Consortium, 1995). PKD2 is homologous to calcium-regulated cation channels, suggesting that PKD2 also is a cation channel (Chen et al., 1999). Further work will be necessary to determine if PKD1 or PKD2 are ever found on mammalian primary cilia.

TPR Repeats in IFT88 and Tg737 Suggest That These Proteins Are Involved in Protein-Protein Interactions

TPR repeats are degenerate 34-amino acid motifs (Lamb et al., 1995) that are present in tandem arrays in proteins. These arrays are predicted to form super-helices (Hirano et al., 1990) with amphipathic grooves responsible for binding specific target proteins (Das et al., 1998). TPR domains have been found to mediate multiple simultaneous protein-protein interactions in such multiprotein complexes as molecular chaperones and the anaphase-promoting complex (reviewed in Blatch and Lassle, 1999). In IFT88 and Tg737, there are three closely spaced TPR repeats in the amino-terminal half of the protein and another seven TPR repeats in the carboxyl-terminal half of the protein. These two separate TPR domains can serve to bind simultaneously to two separate target proteins. These target proteins can be axonemal subunits that are transported via IFT to the flagellar tip where they are assembled (Piperno et al., 1996). The targets also could be membrane proteins such as receptors and channels, as IFT particles are tightly associated with the flagellar membrane (Kozminski et al., 1995; Pazour et al., 1998). Alternatively, IFT88 c be binding to other subunits of the IFT particle and holding it together. IFT57 is likely to be an interacting protein because it is destabilized in the absence of IFT88.

IFT Is a Conserved Mechanism in the Assembly and Maintenance of Cilia and Flagella

A strong body of evidence indicates that IFT is necessary for assembly and maintenance of all eukaryotic motile and sensory cilia. Previous work has shown that the anterograde motor, kinesin-II, is necessary for assembly and maintenance of cilia and flagella in diverse organisms that include green algae, ciliated protozoa, nematodes, echinoderms, and vertebrates (reviewed in Cole, 1999; Marszalek and Goldstein, 2000). The retrograde motor, cytoplasmic dynein DHC1b/DHC2, also has been shown to be required for assembly of *Chlamydomonas* flagella

(Pazour et al., 1999; Porter et al., 1999) and *Caenorhabditis* sensory cilia (Signor et al., 1999; Wicks et al., 2000). Our initial report on the composition of the *Chlamydomonas* IFT particle proteins showed that IFT52 was homologous to *C. elegans* OSM-6 and that IFT172 was homologous to *C. elegans* OSM-1 (Cole et al., 1998). OSM-1 and OSM-6 are required for assembly of sensory cilia in worms (Collet et al., 1998; Perkins et al., *Dev. Biol.*, 117:456-487, 1986). The involvement of these two nematode proteins in IFT was recently confirmed when GFP-labeled OSM-6 and OSM-1 were both shown to undergo IFT in transformed *C. elegans* (Orozco et al., 1999; Signor et al., 1999). The work in this paper shows that the IFT particle protein IFT88 is required for ciliary assembly in *Chlamydomonas* and that the IFT88 homologue, Tg737, is required for assembly of primary cilia in the kidney of mice. Thus, evidence from a diverse group of eukaryotes shows that both the IFT motors and the IFT particle proteins are required for assembly and maintenance of cilia and flagella. This indicates that IFT is an ancient and conserved mechanism by which eukaryotic cilia and flagella are built and maintained.

IFT Is Likely To Play Important Roles in Many Disease States

In addition to PKD discussed above, there are many other diseases that involve IFT. This includes retinitis pigmentosa (RP), which is a genetic disorder that causes destruction of photoreceptor cells resulting in progressive vision loss. Transport of opsin and other components of the rod outer segment is very important in photoreceptor cells, as ~10% of the outer segment is turned over each day. Transport from the inner segment to the outer segment occurs through the connecting cilium (reviewed in Besharse and Horst, 1990). Kinesin-II and several IFT particle proteins are located in the connecting cilium of photoreceptor cells (Beech et al., 1996, Whitehead et al., 1999; Pazour et al., submitted). Moreover, Marszalek et al. (2000) recently showed that photoreceptor cells lacking kinesin-II accumulate opsin and arrestin in the inner segment, indicating that kinesin-II is involved in transport in photoreceptor cells. Therefore IFT is likely to be an important transport mechanism in vertebrate photoreceptor cells, and mutations in IFT particle proteins are likely to cause vision defects.

Defects in IFT also are likely to affect motile cilia and flagella and be a cause of primary ciliary dyskinesia (PCD). PCD is a syndrome caused by defects in motile cilia and is characterized by male infertility, respiratory disease and *situs inversus*. It is well known that defects in axonemal components cause PCD (Afzelius, 1979; Pennarun et al., 1999). However,

IFT particle proteins are highly expressed in the testis and lung, suggesting that they are involved in the assembly of motile sperm flagella and respiratory tract cilia (Pazour, unpublished). Thus, a mutation in an IFT particle protein can lead to defects in sperm flagellar assembly and result in sperm with short disorganized tails as have been described in some infertile human males (Baccetti et al., 1993). It is possible that mutations that prevent assembly of both motile and sensory cilia are so severe that the embryos terminate during gestation, as has been observed with mutations in the mouse kinesin-II motor subunits *kif*3a (Marszalek et al., 1999; Takeda et al.,1999) and *kif*3b (Nonaka et al., 1998).

Example 5. IFT Proteins in Vertebrate Photoreceptor Connecting Cilium

Introduction

Vertebrate photoreceptors are polarized sensory neurons consisting of a photosensitive outer segment, and an inner segment that supports synthesis of proteins destined for all cellular compartments. Transport of the visual pigment, opsin, and other components of the phototransduction machinery from the inner segment to the outer segment is very important in photoreceptor cells, as ~10% of the outer segment is turned over each day (Young, *J. Cell Biol.* 33:61-72 (1967). Although disruption of this process of intersegmental transport results in photoreceptor degeneration and blindness, the mechanisms involved have not been identified. Outer segments develop from a primary cilium which remains as the sole connecting link and presumably the major transport corridor between inner and outer segments (Besharse and Horst, In *Ciliary and Flagellar Membranes* (ed Bloodgood, R.A.) pp. 389-417, Plenum Publishing Corp, New York, 1990). Recently, a process called intraflagellar transport (IFT) has been found to be essential for the assembly and maintenance of sensory cilia in *Caenorhabditis* and the motile cilia of the green alga, *Chlamydomonas*, and sea urchins (Kozminski et al., *J. Cell Biol.* 131:1517-1527, 1995; Collet et al., *Genetics* 148:187-200 1998; Morris et al., *J. Cell Biol.*, 138:1009-1022, 1997). During IFT, kinesin-II transports a large protein complex from the cell body to the tip of the flagellum (Piperno and Mead, *Proc. Natl. Acad. Sci. USA*, 94:4457-4462, 1997; Cole et al., *J. Cell Biol.*, 141:993-1008, 1998; Orozco et al., *Nature* 398:674, 1999; Signor et al, *Mol. Biol. Cell* 10:345-360, 1999). These particles are then returned to the cell body by the DHC1b/DHC2 form of cytoplasmic dynein (Pazour et al., *J. Cell Biol.,* 141:979-992, 1998; Pazour et al., *J. Cell Biol.,* 144:473-481, 1999; Porter et al., *Mol. Biol. Cell,* 10:693-712,

1999; Signor et al., *J. Cell Biol.*, 147:519-530, 1999).). IFT particles are composed of ~17 proteins and are thought to carry components needed for assembly and maintenance of cilia and flagella. They also transport signals between cilia and the cell body (Rosenbaum et al., *J. Cell Biol.*, 144:385-388, 1999). The functions of individual IFT particle proteins are not known but mutations in genes encoding IFT particle proteins in *Chlamydomonas* and *Caenorhabditis* prevent ciliary assembly (Pazour et al., manuscript in preparation). IFT particle proteins are conserved among green algae, nematodes, and vertebrates. To investigate the possibility that IFT functions in sensory cilia of vertebrates, we have identified mouse homologues of three *Chlamydomonas* IFT particle proteins, have generated antibodies against the mouse proteins, and localized the proteins within the retina.

Cloning IFT20: *Chlamydomonas* IFT20 was purified and the sequence of two tryptic peptides was obtained (GVYFDEDFHVR and YVSAIDQQVER) (Cole et al., *J. Cell Biol.* 141:993-1008, 1998). A degenerate PCR primer designed from the first peptide sequence was used in combination with an oligo-dT primer to amplify most of the coding sequence from reverse-transcribed cDNA. The remainder of the gene was amplified from a *Chlamydomonas* cDNA library in lambda ZapII (Stratagene) with a vector primer (M13Rev) and a IFT20-specific primer designed from the sequence of the first PCR product. The open reading frame contained within these clones encodes a 15.6-kD peptide containing both tryptic peptides.

Bioinformatics: Mouse and human homologues of the *Chlamydomonas* IFT particle proteins were identified by BLAST searches of Genbank. The human genetic map positions were determined using publicly available data by the following method. First, a mapped sequence tagged site (STS) from the gene itself (e.g. IFT20-1 and IFT52) or encoded within a genomic clone that also encodes the IFT gene was identified in the STS portion of Genbank. This allowed all the genes to be placed on the human radiation hybrid map. The approximate cytogenetic positions were then predicted using the data of Bray-Ward et al. (Bray-Ward et al., *Genomics* 15:1-14, 1996)..

Antibody production: The antibodies were produced in rabbits by injecting bacterially expressed maltose-binding fusion proteins made by cloning the open reading frames of mouse NDG5 (IFT52), IFT57, and IFT20 into the pMalc expression vector (New England Biolabs,

MA). The sera were affinity purified using immobilized glutathione S-transferase fusion proteins. The latter fusion proteins were made by cloning the open reading frames of the IFT genes into the pGEX-6p-1 (Amersham Pharmacia Biotech, NJ) expression vector.

Preparation of the detergent extracted photoreceptor cytoskeleton (DEPC) and western blotting: The DEPC was prepared as described in Horst et al (Horst et al., *J. Cell Biol.* 105:2973-2987, 1987). Briefly, rod outer segments shaken from 50 dark-adapted, frozen bovine retinas in Buffer A (10mM PIPES, pH 7.0, 1mM EDTA, 5mM $MgCl_2$, 0.02% NaN_3), supplemented with a protease inhibitor cocktail (1ug/ml pepstatin, 1ug/ml leupeptin, 4 ug/ml aprotinin, 1mM benamidine, 1mM PMSF) were purified by sucrose density centrifugation. Outer segments were then extracted in Buffer A containing 2% Triton X-100. Ciliary axonemes were separated from detergent-soluble material by centrifugation over a 45%-60% linear sucrose gradient. Equal aliquots were separated on a 12% denaturing polyacrylamide gel and transferred to Immobilon membranes for antibody labeling and detection using the SuperSignal West Femto chemiluminescent system (Pierce Chemical Co., IL).

Immunocytochemistry: Fresh mouse or bovine retinal tissue was placed in Tissue Freezing Medium™ (Triangle Biomedical Sciences, NC) and quickly frozen in liquid nitrogen with or without prior fixation in 4% paraformadehyde; *Xenopus* retina was prepared in the same way except that fixation was in cold methanol (Whitehead et al., *Exp. Eye Res.* 69:491-503, 1999). Fish inner segment-outer segment preparations were made using procedures described in Beech et al. (Beech et al., *J. Cell Sci.*,109:889-897, 1996). Primary antibodies were detected with goat anti-rabbit or goat anti-mouse IgG conjugated with Cy3 (Jackson Laboratories), Texas Red™, fluoroscein, Alexa 488, or Alexa 568 (Molecular Probes, Eugene OR). In double-label experiments, discrimination of signals for K26 versus IFT proteins involved use of conjugated anti-mouse and anti-rabbit antibodies respectively. For discrimination of two monoclonal antibodies (K26 versus tubulin or opsin), we labeled with one monoclonal antibody and a fluorescent anti-mouse antibody, and then repeated the procedure for the second monoclonal antibody using a different fluorophore. Images of cells labeled with more than one fluorophore were pseudocolored and merged using NIH Image 1.62 or Adobe Photoshop.

Identification of vertebrate homologues of *Chlamydomonas* IFT particle proteins

The *Chlamydomonas* IFT20, IFT52, and IFT57 particle proteins were purified, partially sequenced, and the peptides compared to sequences in GenBank. Three peptides from IFT52 closely matched a rodent protein of ~52 kDa called NGD5, and a *C. elegans* protein called OSM-6. The function of NGD5 is unknown, but its expression is down-regulated by exposing cultured cells to opioids (Wick et al., *Mol. Brain Res.*, 32:171-175, 1995). OSM-6 is required for assembly of sensory cilia in nematodes. The sequence of the *Chlamydomonas* peptides from IFT57 and IFT20 did not show strong similarity to any proteins in the databases. Consequently, degenerate PCR was used to clone these two genes from *Chlamydomonas*. The full-length cDNA sequence of *Chlamydomonas* IFT57 strongly matched a number of human and mouse EST sequences. To determine the extent of homology between the *Chlamydomonas* and mouse proteins, a mouse IFT57 cDNA clone was completely sequenced. The *Chlamydomonas* and mouse proteins are 38% identical and have a BLAST p-value of 3e-67, indicating that the proteins are very likely to have related functions (Pazour et al., manuscript in preparation). The *Chlamydomonas* IFT20 sequence strongly matched a small protein in mouse as well as EST sequences from humans, cattle and a large number of other vertebrate and invertebrate species. The *Chlamydomonas* and mouse proteins are 32% identical and have a BLAST p-value of 4e-15. Bovine, mouse and human IFT20 proteins are nearly identical to each other.

We have obtained data indicating the chromosome locations of the three human IFT protein homologues . Although none map to a region currently associated with retinal degeneration[s], these data will prove useful in future studies of photoreceptor disease or diseases resulting from defects in ciliary development. Analysis of genomic sequence data indicated that in addition to the IFT20 sequence (IFT20-1) most similar to *Chlamydomonas* IFT20, at least two additional related genes are present in man; all three map to different chromosomes. All but one of the IFT20 EST sequences identified corresponded to IFT20-1. The EST corresponding to IFT20-2 has a stop codon within the region of homology, suggesting that it contains a sequencing error or does not encode a functional protein.

Localization of IFT particle proteins to the vertebrate photoreceptor:

Initial evidence for association of IFT particle proteins with photoreceptor cilia came from western blot analysis. Bovine retinal tissue was used for this analysis in order to take advantage of a connecting cilium-specific monoclonal antibody (K26) for immunocytochemistry (Horst et al., *Cell Motil. Cytoskeleton* 17:329-344, 1990) and a procedure for production of a

detergent-extracted photoreceptor cytoskeletal (DEPC) fraction from bovine retina (Horst et al., *J. Cell Biol.* 105:2973-2987, 1987). Affinity-purified antibodies to IFT20, IFT52 and IFT57 did not readily detect proteins in whole cell extracts of bovine retina, but strongly detected single bands of ~16, 52 and 57 kDa respectively in the DEPC fraction. Because this fraction is highly enriched in ciliary axonemes from photoreceptors, it seemed very likely that the IFT proteins were associated with the photoreceptor cilia.

Immunocytochemical analysis revealed that IFT20, IFT52, and IFT57 were most abundant in the inner segments (IS) of mouse (not shown) and bovine photoreceptors, with distinctly less labeling over the outer segments and other regions of the photoreceptor cells. The signal in the inner segment was distinctly granular in appearance, particularly at the junction between the inner and outer segments where the connecting cilia are located. The outer nuclear layer (ONL) containing photoreceptor nuclei exhibited perinuclear staining. The inner nuclear layer (INL) was also labeled; the latter was most easily seen with antibodies to IFT52. All three antibodies labeled the outer plexiform layer which contains the synaptic terminals of photoreceptors. This and the presence of kinesin II in these synapses (Muresan et al., *J. Neurosci.*, 19:1027-1037, 1999) suggests that IFT proteins have functions in photoreceptor synaptic terminals.

Double-label immunocytochemistry with a monoclonal antibody (K26) that recognizes a connecting cilium specific epitope demonstrated that all three IFT proteins are associated with the ciliary axoneme in situ. The K26 antibody uniquely stained the connecting cilium at the base of photoreceptor outer segments, which were identified with antibodies to rod opsin. In contrast, antibodies to acetylated alpha-tubulin labeled microtubules of the inner segment and the ciliary axoneme. Labeling by both K26 (red) and acetylated alpha tubulin (green) antibodies in the connecting cilium resulted in a yellow to orange color due to overlap in the connecting cilium, and demonstrated that axonemal microtubules extend distally beyond the connecting cilium into the outer segment. Antibodies to IFT20, IFT52 and IFT57 labeled structures both on the proximal (inner segment) and distal (outer segment) side of the connecting cilium in a large proportion of the photoreceptors. Frequent yellow to orange coloration of the connecting cilium was indicative of overlap of the two labels (red, K26 and green, anti-IFT antibody) within the connecting cilium. Triple-labeled images revealed a similar pattern in which IFT57 (blue) is

found in association with microtubules (acetylated α-tubulin) at both ends of the connecting cilium (K26) in most but not all photoreceptors.

Association of IFT proteins with ciliary axonemes was also detected in the large rod photoreceptors of the frog, *Xenopus laevis* , and the fish, *Lepomis macrochirus*. Confocal analysis of *Xenopus* tissue sections double labeled with antibodies to acetylated α-tubulin and IFT57 revealed a single area of intense immunoreactivity in each cell corresponding to the base of the ciliary axoneme, in the region of the basal body. A similar pattern was seen in whole-mounted, isolated cells of *Lepomis*, although the zone of staining at the base of the axoneme was broader than in *Xenopus*. IFT57 staining of the distal axoneme within the outer segment was sometimes seen as punctate spots in both Xenopus and Lepomis. These staining patterns are similar to those seen in *Chlamydomonas*. In *Chlamydomonas*, the IFT particle proteins are located primarily in a cytoplasmic pool at the base of the cilia with only a small number of punctate spots found along the length of the cilia. The punctate staining is thought to be due to IFT particles that were in transit when the cells were fixed.

Discussion:

Macromolecules of the photoreceptor outer segment are synthesized in the inner segment and transported into the outer segment. This process occurs at a prodigious rate. It has been estimated that turnover in each mammalian photoreceptor outer segment requires delivery of as many as 2000 photopigment molecules per minute throughout the life of the cell. In the larger photoreceptors of amphibians, this rate is increased by more than an order of magnitude. In addition to the photopigment molecules, proteins of the phototransduction machinery (Philp et al., *FEBS Lett.*, 225:127-132, 1987; Whelan et al., *J. Neurosci. Res.* 20:263-270, 1988) and phospholipid components of the discs (Anderson et al., *Biochim. Biophys. Acta,* 620:212-226, 1980) turn over rapidly. Although transport between the inner and outer segment is crucial to polarized organization of the cell, the underlying mechanism has remained elusive.

The presence of IFT proteins in photoreceptor cilia strongly suggests that IFT is an important transport mechanism in these cells. Although the transported cargo has not been identified, our data allow us to propose the following model based on the idea that membrane components, including rhodopsin and phospholipid, and many soluble proteins are normally targeted to the photoreceptor outer segment. Membrane proteins are synthesized on the endoplasmic reticulum and modified during passage through the Golgi network. The polarity of

inner segment microtubules with their minus ends associated with the base of the cilium suggests that these membrane proteins and phospholipids are transported from the Golgi stack to the base of the connecting cilium by dynein complexes containing the DHC1 heavy chain. DHC1 is a well established vesicle transporter and has been shown to interact with the cytoplasmic tail of rhodopsin (Tai et al., *Cell* 97:877-887, 1999). Cytoplasmic proteins destined for the outer segment such as components of the transduction machinery (transducin and arrestin) and the ciliary axoneme are also synthesized in the inner segment and can be transported in association with IFT particles. At the base of the connecting cilium where IFT proteins are normally most concentrated, they associate with the surface of these vesicles and with other outer segment proteins. Once the vesicles fuse with the plasma membrane adjacent to the cilium, transport of the complex with attached cargo along the connecting cilium would occur by the kinesin-II motor. Kinesin-II is thought to be the anterograde IFT motor in *Chlamydomonas* and *Caenorhabditis* and has been localized to the connecting cilium in vertebrates (Beech et al., *J. Cell Sci.*, 109:889-897, 1996; Whitehead et al., *Exp. Eye Res.*, 69:491-503, 1999). In the outer segment, the IFT particles disengage from their cargo, and the membrane is organized into disks and pinched off by myosin VIIA. Myosin VII is required for phagocytosis in Dictyostelium (Titus, *Curr Biol.*, 9:1297-1303, 1999), and mice with defects in the myosin VIIA gene accumulate opsin in the connecting cilium (Liu et al., *J. Neurosci.* 19:6267-6274, 1999). Soluble proteins of the transduction machinery and cytoskeleton would be expected to associate with appropriate protein complexes within the outer segment, while IFT particles at the distal end of the connecting cilium would be returned to the base of the cilium by a dynein complex containing the DHC1b/DHC2 heavy chain. DHC1b/DHC2 is the retrograde IFT motor in *Chlamydomonas*[10-12] and *Caenorhabditis* (Signor, *J. Cell Biol.*, 147:519-530, 1999) and has been localized to the connecting cilium of vertebrate photoreceptors (Besharse et al., in preparation). At the base of the connecting cilium the IFT particles re-enter a peri-basal body pool of IFT particle proteins and begin the cycle again. The IFT particles also move components of the transduction machinery from the outer segment to the inner segment. For example, both transducin and arrestin have been shown to rapidly move between the segments during light and dark adaptation.

In conclusion, we have shown that three different IFT particle proteins are localized to the connecting cilium of vertebrate photoreceptor cells. Since IFT is essential for assembly and

maintenance of motile and sensory cilia of *Chlamydomonas* and *Caenorhabditis*, it is likely that it also is important in vertebrate photoreceptors. The availability of mouse mutations in the genes encoding kinesin-II (Nonaka, *Cell* 95:829-837, 1998; Bray-Ward et al., *Genomics*, 15:1-14, 1996) and an IFT particle protein (Pazour et al., manuscript in preparation) will allow us to directly test this hypothesis.

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

LITERATURE CITED

Afzelius et al., Immotile-cilia syndrome (primary ciliary dyskinesia), including Kargagener syndrome. In: The Metabolic and Molecular Bases of Inherited Disease, Vol. III. Ed: Scriver, C. R. et al. McGraw-Hill, Inc., NY. pp. 3943-3954 (1995).

Alberts, B. et al. Molecular Biology of the Cell. Third Ed. Garland Publishing Inc. (1994) p. 820.

Barr et al., *Nature* **401**, 386-389 (1999).

Beech et al., *J. Cell Sci.* **109**, 889-897 (1996).

Bergman et al., *J. Cell Biol.* **3,** 606-622 (1975).

Besharse, J. C. Photosensitive membrane turnover: Differentiated membrane domains and cell-cell interaction, in: *The Retina: A Model for Cell Biological Studies,* Part 1 (R. Adler and D. Farber, eds.), Academic Press, New York, pp. 297-352 (1986).

Bloodgood, Exp. Cell Res. **150,** 488-493 (1984).

Bloodgood, *Cell Biol. Int.* **24,** 857-862 (2000).

Bouck, *J. Cell Biol.* **50,** 362-384 (1971).

Blyth et al., *J. Med. Genet.* **8**, 257-284 (1971).

Cole et al., *J. Pediatr.* **111,** 693-699 (1987).

Cole, D. G. et al., *J. Cell Biol.* **141**, 993-1008 (1998).

Collet et al., *Genetics* **148**, 187-200 (1998).

Dentler, W. L. Linkages between microtubules and membranes in cilia and flagella. In: Ciliary and Flagellar Membranes. Ed: Bloodgood, R. A. Plenum Press, NY. pp. 31-64 (1990).

Dentler et al., *J. Cell Biol.* **74,** 747-759 (1977).

De Robertis, E., *J. Biophys. Biochem. Cytol. Suppl.* **2,** 209-216.

Diener et al., *Mol. Biol. Cell* **7,** 47a (1996).

Emmons et al., *Nature* **401,** 339-340 (1999).

Espindola et al., *Cell Motil. Cytoskeleton* **47,** 269-281 (2000).

Fowkes et al., *Mol. Biol. Cell* **9,** 2337-2347 (1998).

Godsel et al., *EMBO Journal* **18,** 2057-2065 (1999).
Gonzalez-Perret, S., et al., *Proc. Natl. Acad. Sci. U.S.A.* **98,** 1182-1187 (2001).
Handel, M., et al., *Neuroscience* **89,** 909-926 (1999).
Johnson et al., *J. Cell Biol.* **119,** 1605-1611 (1992).
King, S. M., et al., *J. Biol. Chem.* **271,** 19358-19366 (1996).
Kozminski et al., *Proc. Nat. Acad. Sci. U.S.A.* **90**, 5519-5523 (1993).
Kozminski et al., *J. Cell Biol.* **131**, 1517-1527 (1995).
Marshall, W. F. & Rosenbaum, J. L., *Mol. Biol. Cell* **11,** 368a. (2000c).
Marszalek et al., *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5043-5048 (1999).
Marszalek, J. R. et al., *Cell* **102**, 175-187 (2000).
Morris, R. L. & Scholey, J. M., *J. Cell Biol.* **138**, 1009-1022 (1997).
Moyer, J. H., et al., *Science* **264,**1329-1333 (1994).
Murcia et al., *Kidney Intl.* **55,** 1187-1197 (1999).
Orozco et al., *Nature* **398**, 674 (1999).
Pan, J. & Snell, W. J., *Curr. Opin. Microbiol.* **6,** 596-602 (2000b).
Pan, J. & Snell, W. J., *Mol. Biol. Cell* **11,** 368a (2000c).
Pazour et al., *J. Cell Biol.* **141**, 979-992 (1998).
Pazour et al., *J. Cell Biol.* **144**, 473-481 (1999).
Pazour et al., *Mol. Biol. Cell* **10,** 369a (1999b).
Pazour et al., *Mol. Biol. Cell* **10,** 388a. (1999c).
Pazour, G. J., et al., *J. Cell Biol.* **151,** 709-718 (2000).
Pazour, G. J., et al., *Mol. Biol. Cell* **11,** 540a (2000b).
Perkins et al., *Dev. Biol.* **117,** 456-487 (1986).
Piperno et al., *J. Cell Biol.* **133**, 371-379 (1996).
Piperno et al., *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4457-4462 (1997).
Piperno et al., *J. Cell Biol.* **143,** 1591-1601 (1998).
Qin et al., Cur. Bio. **11,** 1-20 (2001).
Scholey, J. M., *J. Cell Biol.* **133,** 1-4 (1996).
Signor et al., *J. Cell Biol.* **147**, 519-530 (1999).
Snapp et al., *J. Cell Biol.* **139,** 1775-1783 (1997).
Snapp et al., *J. Biol Chem.* **274,** 29543-29548 (1999).
Somlo, S. & Ehrlich, B., *Current Biol.* **11,** R356-R360 (2001).
Sung, C-H. & Tai, A.W., *Internat. Rev. Cytology* **195,** 0074-7696 (2000).
Supp et al., *Trends Cell Biol.* **10,** 41-45 (2000).
Wheatley et al., *Cell Biol. Int.* **20**, 73-81 (1996).
Wicks et al., *Dev. Biol.* **221,** 295-307 (2000).
Young, R. W. *Invest. Opthalmol.* **15,** 700-725 (1976).
Afzelius, B.A. 1979, *Int. Rev. Exp. Pathol.* 19:1-43.
Baccetti et al., 1993, *Andrologia* 25:331-335.
Cole, D.G. 1999. Kinesin-II, coming and going. *J. Cell Biol.* 147:463-466.
Flood, P.R. and G.K. Totland. 1977, *Cell Tiss. Res.* 183:281-290.
Hughes, et al., 1995, *Nat. Genet.* 10:151-160.
Motta et al., 1974, *Z. Anat .Entwickl.-Gesch.* 145:223-226.
Pazour et al., 1999, *Mol. Biol. Cell* 10:3507-3520.
Pazour et al., 1995, *J. Cell Biol.* 131:427-440.
Richards et al., 1997, *Am. J. Path.* 150:1189-1197.
Rosenbaum et al., 1999, *J. Cell Biol.* 144:385-388.